

# Invariant higher-order variational problems

François Gay-Balmaz<sup>1</sup>, Darryl D. Holm<sup>2</sup>, David M. Meier<sup>2</sup>,  
Tudor S. Ratiu<sup>3</sup>, François-Xavier Vialard<sup>2</sup>

*Fondly remembering our late friend Jerry Marsden*

**AMS Classification:**

**Keywords:**

## Abstract

We investigate higher-order geometric  $k$ -splines for template matching on Lie groups. This is motivated by the need to apply diffeomorphic template matching to a series of images, e.g., in longitudinal studies of Computational Anatomy. Our approach formulates Euler-Poincaré theory in higher-order tangent spaces on Lie groups. In particular, we develop the Euler-Poincaré formalism for higher-order variational problems that are invariant under Lie group transformations. The theory is then applied to higher-order template matching and the corresponding curves on the Lie group of transformations are shown to satisfy higher-order Euler-Poincaré equations. The example of  $SO(3)$  for template matching on the sphere is presented explicitly. Various cotangent bundle momentum maps emerge naturally that help organize the formulas. We also present Hamiltonian and Hamilton-Ostrogradsky Lie-Poisson formulations of the higher-order Euler-Poincaré theory for applications on the Hamiltonian side.

---

<sup>1</sup> Laboratoire de Météorologie Dynamique, École Normale Supérieure/CNRS, Paris, France. [gaybalma@lmd.ens.fr](mailto:gaybalma@lmd.ens.fr)

<sup>2</sup>Department of Mathematics and Institute for Mathematical Sciences, Imperial College, London SW7 2AZ, UK. [d.holm@ic.ac.uk](mailto:d.holm@ic.ac.uk), [d.meier09@ic.ac.uk](mailto:d.meier09@ic.ac.uk), [f.vialard@ic.ac.uk](mailto:f.vialard@ic.ac.uk)

<sup>3</sup>Section de Mathématiques and Bernoulli Center, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. [tudor.ratiu@epfl.ch](mailto:tudor.ratiu@epfl.ch)

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Previous work on geometric splines for trajectory planning and interpolation . . . .	3
1.2	Main content of the paper . . . . .	4
<b>2</b>	<b>Geometric setting</b>	<b>6</b>
2.1	$k^{\text{th}}$ -order tangent bundles . . . . .	6
2.2	$k^{\text{th}}$ -order Euler-Lagrange equations . . . . .	7
2.3	Quotient space and reduced Lagrangian . . . . .	8
<b>3</b>	<b>Higher-order Euler-Poincaré reduction</b>	<b>10</b>
3.1	Quotient map, variations and $k^{\text{th}}$ -order Euler-Poincaré equations . . . . .	10
3.2	Example: Riemannian cubics . . . . .	12
3.3	Parameter dependent Lagrangians . . . . .	17
3.4	Splines with constraints . . . . .	19
<b>4</b>	<b>Clebsch-Pontryagin optimal control</b>	<b>21</b>
<b>5</b>	<b>Higher-order template matching problems</b>	<b>23</b>
5.1	Previous work on longitudinal data interpolation in CA . . . . .	23
5.2	Euler-Lagrange equations for higher-order template matching . . . . .	23
5.3	Two examples of interest for computational anatomy . . . . .	27
5.4	Template matching on the sphere . . . . .	27
<b>6</b>	<b>Optimization with penalty</b>	<b>30</b>
<b>7</b>	<b>Clebsch and Lie-Poisson-Ostrogradsky formulations</b>	<b>33</b>
7.1	Higher order Clebsch formulations . . . . .	34
7.2	Ostrogradsky-Lie-Poisson reduction . . . . .	37
<b>8</b>	<b>Outlook and open problems</b>	<b>41</b>
8.1	Brief summary and other potential directions . . . . .	41
8.2	An open problem: the slalom, or brachistochrone for splines . . . . .	42

## 1 Introduction

**The purpose of this paper.** This paper provides a method for taking advantage of continuous symmetries in solving Lie group invariant optimization problems for cost functions that are defined on  $k^{\text{th}}$ -order tangent spaces of Lie groups. The type of application we have in mind is, for example, the interpolation and comparison of a series of images in longitudinal studies in a biomedical setting.

Previous work on the geometric theory of Lagrangian reduction by symmetry on first-order tangent spaces of Lie groups provides a convenient departure point that is generalized here to allow for invariant variational problems formulated on higher-order tangent spaces of Lie groups. It turns out that this generalization may be accomplished as a series of adaptations of previous advances in Euler-Poincaré theory, placed into the context of higher-order tangent spaces. Extension of the basic theory presented here to allow for actions of Lie groups on Riemannian manifolds should have several interesting applications, particularly in image registration, but perhaps elsewhere, too. Actions of Lie groups on Riemannian manifolds will be investigated in a subsequent treatment. Two important references for the present work are [HMR98] for the basic Euler-Poincaré theory and [CMR01] for the bundle setting of geometric mechanics.

## 1.1 Previous work on geometric splines for trajectory planning and interpolation

The topics treated here fit into a class of problems in control theory called *trajectory planning and interpolation by variational curves*. These problems arise in numerous applications in which velocities, accelerations, and sometimes higher-order derivatives of the interpolation path need to be optimized simultaneously. Trajectory planning using variational curves in Lie groups acting on Riemannian manifolds has been discussed extensively in the literature. For example, trajectory planning for rigid body motion involves interpolation on either the orthogonal group  $SO(3)$  of rotations in  $\mathbb{R}^3$ , or the semidirect-product group  $SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$  of three-dimensional rotations and translations in Euclidean space. Trajectory planning problems have historically found great utility with applications, for example, in aeronautics, robotics, biomechanics, and air traffic control.

Investigations of the trajectory planning problem motivated the introduction in [GK85] and [NHP89] of a class of variational curves called *Riemannian cubics*. Riemannian cubics and their recent higher order generalizations are reviewed in [Pop07] and [MSK10], to which we refer for extensive references and historical discussions. The latter work addresses the interpolation by variational curves that generalizes the classical least squares problem to Riemannian manifolds. This generalization is also based on the formulation of higher-order variational problems, whose solutions are smooth curves minimizing the  $L^2$ -norm of the covariant derivative of order  $k \geq 1$ , that fit a given data set of points at given times. These solutions are called  $k^{\text{th}}$ -order geometric splines, or geometric  $k$ -splines. This approach was initiated in [NHP89] for the construction of smoothing splines with  $k = 2$  for the Lie group  $SO(3)$  and then generalized to higher order in [CSC95]. The following result, noted in the first of these papers and then discussed more generally in the second one, was another source of motivation for the present work.

**Proposition 1.1** ([NHP89]).

*The equation for a  $2^{\text{nd}}$ -order geometric spline for a bi-invariant metric on  $SO(3)$  may be written as a dynamical equation for a time-dependent vector  $\Omega(t) \in \mathbb{R}^3$  using the vector cross product*

$$\ddot{\Omega} = \dot{\Omega} \times \Omega, \tag{1.1}$$

*for all  $t$  in a certain interval  $[0, T]$ .*

Solutions of the more general version of equation (1.1) expressed in [CS95] for  $2^{\text{nd}}$ -order geometric splines on Lie groups in terms of the Lie algebra commutator are called ‘Lie quadratics’ in [Noa03, Noa04, Noa06].

As we said, understanding the intriguing result in Proposition 1.1 from the viewpoint of Lie group-invariant higher-order variational principles was one of the motivations for the present work. Its general version is proved again below as equation (3.21) in Section 3 by using the Euler-Poincaré methods of [HMR98] for higher-order variational principles that are invariant under the action of a Lie group. The directness and simplicity of the present proof of the general version of Proposition 1.1 compared with other proofs available in the literature encouraged us to continue investigating the application of Lie group-invariant  $k^{\text{th}}$ -order variational principles for geometric  $k$ -splines. It turns out that higher-order Euler-Poincaré theory is the perfect tool for studying geometric  $k$ -splines.

The Euler-Poincaré theory for first-order invariant variational principles focuses on the study of geodesics on Lie groups, which turns out to be the fundamental basis for both ideal fluid dynamics and modern large-deformation image registration. For reviews and references to earlier work on first-order invariant variational principles, see [HMR98] for ideal fluids and [You10] for large-deformation image registration. The present paper begins by extending these earlier results for geodesics governed by first-order variational principles that are invariant under a Lie group, so as to include dependence on higher-order tangent spaces of the group (i.e., higher-order time derivatives of curves on the group). This extension is precisely what is needed in designing geometric  $k$ -splines for trajectory planning problems on Lie groups. The essential strategy in making this extension is the application of reduction by symmetry to the Lagrangian before taking variations, as introduced in [HMR98] for continuum dynamics. The equivalence of the result of Lagrangian reduction by symmetry with the results in the literature for Riemannian cubics and  $k^{\text{th}}$ -order geometric splines is shown in Section 3, Proposition 3.3.

This previous work has created the potential for many possible applications. In this paper, we shall concentrate on the application of these ideas in template matching for Computational Anatomy (CA). Although we do not perform explicit image matching here, we demonstrate the higher-order approach to template matching in the finite dimensional case by interpolating a sequence of points on the sphere  $S^2$ , using  $SO(3)$  as the Lie group of transformations.

## 1.2 Main content of the paper

The *main content* of the paper is outlined as follows:

Section 2 discusses the geometric setting for the present investigation of extensions of group-invariant variational principles to higher order. In particular, Section 2 summarizes the definition of higher order tangent bundles and connection-like structures defined on them, mainly by adapting the treatment in [CMR01] for the geometric formulation of Lagrangian reduction.

Section 3 explains the quotient map for higher-order Lagrangian reduction by symmetry and uses it to derive the basic  $k^{\text{th}}$ -order Euler-Poincaré equations. This extends to higher-order the Euler-Poincaré equations derived in [HMR98]. The  $k^{\text{th}}$ -order Euler-Poincaré equations are then applied to derive the equations for geometric  $k$ -splines on a Lie group. After these preliminary developments, there follows a sequence of adaptations of previous advances in Euler-Poincaré theory to higher-order tangent spaces.

Section 4 extends the Clebsch-Pontryagin approach of [GBR10] to develop the  $k^{\text{th}}$ -order Euler-Poincaré equations for potential applications in optimal control. This extension highlights

the role of coadjoint motion for cotangent-lift momentum maps.

Section 5 addresses theoretical and numerical results for our main motivation, longitudinal data interpolation. That is, interpolation through a sequence of data points. After a brief account of the previous work done in Computational Anatomy (CA), we derive the equations that generalize the equations for geodesic template matching [GBHR10] to the case of *higher-order* cost functionals and sequences of *several* data points. We recover in particular the higher-order Euler-Poincaré equations. For a particular choice of cost functionals one can therefore think of the higher-order template matching approach as *template matching by geometric  $k$ -splines*. We discuss the gain in smoothness afforded by the higher-order approach, then we provide a qualitative discussion of two Lagrangians that are of interest for applications in CA. Finally, we close the section by demonstrating the higher-order approach to template matching in the finite dimensional case by interpolating a sequence of points on the sphere  $S^2$ , using  $SO(3)$  as the Lie group of transformations. This yields the template-matching analog of the NHP equation of [NHP89] in (1.1). The results are shown as curves on the sphere in Figures 5.2. A sample figure is shown below to explain the type of results we obtain.

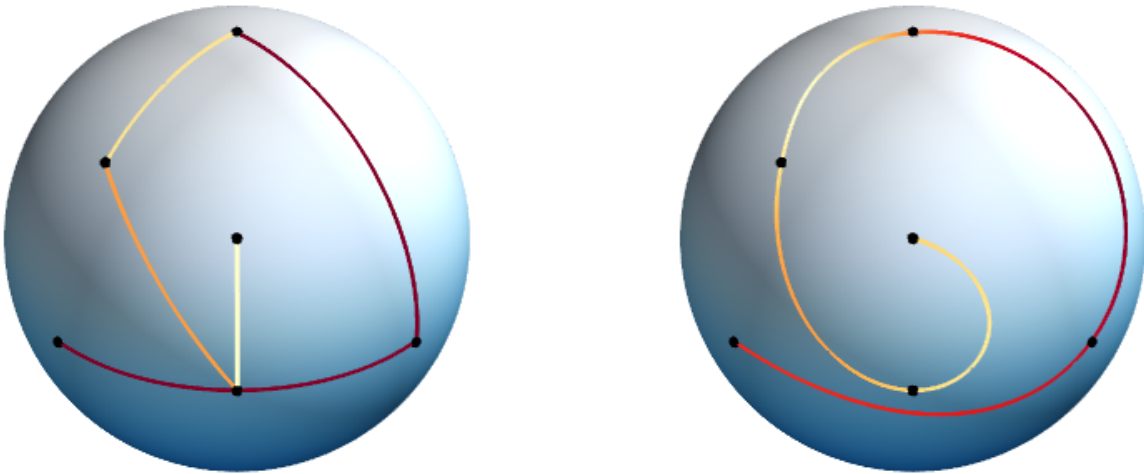


Fig. 1.1: First order *vs.* second order template matching results interpolating a sequence of evenly time-separated points on the sphere, using a bi-invariant metric on the rotation group  $SO(3)$ . The colors show the local speed along the curves on the spheres (white smaller, red larger). The motion slows as the curve tightens.

Section 6 extends to  $k^{th}$ -order tangents the metamorphosis approach of [HTY09] for image registration and the optimization dynamics introduced in [GBHR10].

Section 7 addresses Hamiltonian and Hamilton-Ostrogradsky formulations of the higher-order Euler-Poincaré theory. The Hamilton-Ostrogradsky formulation results in a compound Poisson bracket comprising a sum of canonical and Lie-Poisson brackets.

Section 8 discusses the outlook for future research and other potential applications of the present approach. These include the formulation of higher-order Lie group invariant variational principles that include both curves on Lie groups and the actions of Lie groups on smooth manifolds, and the formulation of a  $k$ th-order brachistochrone problem.

This paper represents only the beginning of our work in this direction. The extensions to higher order discussed here demonstrate the unity and versatility of the geometric approach. We hope these methods will be a source of inspiration for future analysis and applications of Lie group reduction of higher-order invariant variational problems.

## 2 Geometric setting

We shall begin by reviewing the definition of higher order tangent bundles and the connection-like structures defined on them. For more details and explanations of the geometric setting for higher-order variational principles see [CMR01].

### 2.1 $k^{\text{th}}$ -order tangent bundles

The  $k^{\text{th}}$ -order tangent bundle  $\tau_Q^{(k)} : T^{(k)}Q \rightarrow Q$  is defined as the set of equivalence classes of  $C^k$  curves in  $Q$  under the equivalence relation that identifies two given curves  $q_i(t), i = 1, 2$ , if  $q_1(0) = q_2(0) = q_0$  and in any local chart we have  $q_1^{(l)}(0) = q_2^{(l)}(0)$ , for  $l = 1, 2, \dots, k$ , where  $q^{(l)}$  denotes the derivative of order  $l$ . The equivalence class of the curve  $q(t)$  at  $q_0 \in Q$  is denoted  $[q]_{q_0}^{(k)}$ . The projection

$$\tau_Q^{(k)} : T^{(k)}Q \rightarrow Q \quad \text{is given by} \quad \tau_Q^{(k)}([q]_{q_0}^{(k)}) = q_0.$$

It is clear that  $T^{(0)}Q = Q$ ,  $T^{(1)}Q = TQ$ , and that, for  $0 \leq l < k$ , there is a well defined fiber bundle structure

$$\tau_Q^{(l,k)} : T^{(k)}Q \rightarrow T^{(l)}Q, \quad \text{given by} \quad \tau_Q^{(l,k)}([q]_{q_0}^{(k)}) = [q]_{q_0}^{(l)}.$$

Apart from the cases where  $k = 0$  and  $k = 1$ , the bundles  $T^{(k)}Q$  are not vector bundles. The bundle  $T^{(2)}Q$  is often denoted  $\ddot{Q}$ , and is called the **second order bundle**.

**Remark 2.1.** We note that  $T^{(k)}Q = J_0^k(\mathbb{R}, Q)$  consists of  $k$ -jets of curves from  $\mathbb{R}$  to  $Q$  based at  $0 \in \mathbb{R}$ , as defined, for example, in [Bou71, §12.1.2].

A smooth map  $f : M \rightarrow N$  induces a map

$$T^{(k)}f : T^{(k)}M \rightarrow T^{(k)}N \quad \text{given by} \quad T^{(k)}f([q]_{q_0}^{(k)}) := [f \circ q]_{f(q_0)}^{(k)}. \quad (2.1)$$

In particular, a group action  $\Phi : G \times Q \rightarrow Q$  naturally lifts to a group action

$$\Phi^{(k)} : G \times T^{(k)}Q \rightarrow T^{(k)}Q \quad \text{given by} \quad \Phi_g^{(k)}([q]_{q_0}^{(k)}) := T^{(k)}\Phi_g([q]_{q_0}^{(k)}) = [\Phi_g \circ q]_{\Phi_g(q_0)}^{(k)}. \quad (2.2)$$

This action endows  $T^{(k)}Q$  with a principal  $G$ -bundle structure. The quotient  $(T^{(k)}Q)/G$  is a fiber bundle over the base  $Q/G$ . The class of the element  $[q]_{q_0}^{(k)}$  in the quotient  $(T^{(k)}Q)/G$  is denoted  $\left[ [q]_{q_0}^{(k)} \right]_G$ .

**The case of a Lie group.** The  $k^{\text{th}}$ -order tangent bundle  $T^{(k)}G$  of a Lie group  $G$  carries a natural Lie group structure: if  $[g]_{g_0}^{(k)}$ , and  $[h]_{h_0}^{(k)}$  are classes of curves  $g$  and  $h$  in  $G$ , define  $[g]_{g_0}^{(k)}[h]_{h_0}^{(k)} := [gh]_{g_0h_0}^{(k)}$ . The Lie algebra  $T_e T^{(k)}G$  of  $T^{(k)}G$  can be naturally identified, as a vector space, with  $(k+1)\mathfrak{g}$  (that is, the direct sum of  $k+1$  copies of  $\mathfrak{g}$ ) which, therefore, carries a unique Lie algebra structure such that this identification becomes a Lie algebra isomorphism.

## 2.2 $k^{\text{th}}$ -order Euler-Lagrange equations

Consider a Lagrangian  $L : T^{(k)}Q \rightarrow \mathbb{R}$ ,  $L = L(q, \dot{q}, \ddot{q}, \dots, q^{(k)})$ . Then a curve  $q : [t_0, t_1] \rightarrow Q$  is a critical curve of the action

$$\mathcal{J}[q] = \int_{t_0}^{t_1} L(q(t), \dot{q}(t), \dots, q^{(k)}(t)) dt \quad (2.3)$$

among all curves  $q(t) \in Q$  whose first  $(k-1)$  derivatives are fixed at the endpoints:  $q^{(j)}(t_i)$ ,  $i = 0, 1$ ,  $j = 0, \dots, k-1$ , if and only if  $q(t)$  is a solution of the  $k^{\text{th}}$ -order Euler-Lagrange equations

$$\sum_{j=0}^k (-1)^j \frac{d^j}{dt^j} \frac{\partial L}{\partial q^{(j)}} = 0. \quad (2.4)$$

The corresponding variational principle is Hamilton's principle,

$$\delta \int_{t_0}^{t_1} L(q(t), \dot{q}(t), \dots, q^{(k)}(t)) dt = 0.$$

In the  $\delta$ -notation, an infinitesimal variation of the curve  $q(t)$  is denoted by  $\delta q(t)$  and defined by the *variational derivative*,

$$\delta q(t) := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} q(t, \varepsilon), \quad (2.5)$$

where  $q(t, 0) = q(t)$  for all  $t$  for which the curve is defined and  $\frac{\partial^j q}{\partial t^j}(t_i, \varepsilon) = q^{(j)}(t_i)$ , for all  $\varepsilon$ ,  $j = 0, 1, \dots, k-1$ ,  $i = 0, 1$ . Thus  $\delta q^{(j)}(t_0) = 0 = \delta q^{(j)}(t_1)$  for  $j = 0, \dots, k-1$ . Note that the local notation  $L(q, \dot{q}, \dots, q^{(k)})$  used above can be intrinsically written as  $L([q]_q^{(k)})$ .

**Examples: Riemannian cubic polynomials and generalizations.** As originally introduced in [NHP89], Riemannian cubic polynomials generalize Euclidean splines to Riemannian manifolds. Let  $(Q, \gamma)$  be a Riemannian manifold and  $\frac{D}{Dt}$  be the covariant derivative along curves associated with the Levi-Civita connection  $\nabla$  for the metric  $\gamma$ . The Riemannian cubic polynomials are defined as minimizers of the functional  $\mathcal{J}$  in (2.3) for the Lagrangian  $L : T^{(2)}Q \rightarrow \mathbb{R}$  defined by

$$L(q, \dot{q}, \ddot{q}) := \frac{1}{2} \gamma_q \left( \frac{D}{Dt} \dot{q}, \frac{D}{Dt} \dot{q} \right). \quad (2.6)$$

This Lagrangian is well-defined on the second-order tangent bundle since, in coordinates

$$\frac{D}{Dt} \dot{q}^k = \ddot{q}^k + \Gamma_{ij}^k(q) \dot{q}^i \dot{q}^j, \quad (2.7)$$

where  $(\Gamma_{ij}^k(q))_{i,j,k}$  are the Christoffel symbols at point  $q$  of the metric  $\gamma$  in the given basis. These Riemannian cubic polynomials have been generalized to the so-called elastic splines through the following class of Lagrangians

$$L_\tau(q, \dot{q}, \ddot{q}) := \frac{1}{2} \gamma_q \left( \frac{D}{Dt} \dot{q}, \frac{D}{Dt} \dot{q} \right) + \frac{\tau^2}{2} \gamma_q(\dot{q}, \dot{q}), \quad (2.8)$$



where  $\tau$  is a real constant. Another extension are the higher-order Riemannian splines, or geometric  $k$ -splines, where

$$L_k(q, \dot{q}, \dots, q^{(k)}) := \frac{1}{2} \gamma_q \left( \frac{D^{k-1}}{Dt^{k-1}} \dot{q}, \frac{D^{k-1}}{Dt^{k-1}} \dot{q} \right), \quad (2.9)$$

for  $k > 2$ . As for the Riemannian cubic splines,  $L_k$  is well-defined on  $T^{(k)}Q$ . Denoting by  $R$  the curvature tensor defined as  $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$ , the Euler-Lagrange equation for elastic splines ( $k = 2$ ) reads

$$\frac{D^3}{Dt^3} \dot{q}(t) + R \left( \frac{D}{Dt} \dot{q}(t), \dot{q}(t) \right) \dot{q}(t) = \tau^2 \frac{D}{Dt} \dot{q}(t), \quad (2.10)$$

as proven in [NHP89]. For the higher-order Lagrangians  $L_k$ , the Euler-Lagrange equations read [CSC95]

$$\frac{D^{2k-1}}{Dt^{2k-1}} \dot{q}(t) + \sum_{j=2}^k (-1)^j R \left( \frac{D^{2k-j-1}}{Dt^{2k-j-1}} \dot{q}(t), \frac{D^{j-2}}{Dt^{j-2}} \dot{q}(t) \right) \dot{q}(t) = 0. \quad (2.11)$$

These various Lagrangians can be used to interpolate between given configurations on  $T^{(k)}Q$ . The choice of Lagrangian will depend on the application one has in mind. For instance, the following interpolation problem was addressed in [HB04a] and was motivated by applications in space-based interferometric imaging.

**Interpolation problem.** *Given  $N+1$  points  $q_i \in Q$ ,  $i = 0, \dots, N$  and tangent vectors  $v_j \in T_{q_j}Q$ ,  $j = 0, N$ , minimize*

$$\mathcal{J}[q] := \frac{1}{2} \int_{t_0}^{t_N} \left( \gamma_{q(t)} \left( \frac{D}{Dt} \dot{q}(t), \frac{D}{Dt} \dot{q}(t) \right) + \tau^2 \gamma_{q(t)} (\dot{q}(t), \dot{q}(t)) \right) dt, \quad (2.12)$$

*among curves  $t \mapsto q(t) \in Q$  that are  $C^1$  on  $[t_0, t_N]$ , smooth on  $[t_i, t_{i+1}]$ ,  $t_0 \leq t_1 \leq \dots \leq t_N$ , and subject to the interpolation constraints*

$$q(t_i) = q_i, \quad \text{for all } i = 1, \dots, N-1$$

*and the boundary conditions*

$$q(t_0) = q_0, \quad \dot{q}(t_0) = v_0, \quad \text{and} \quad q(t_N) = q_N, \quad \dot{q}(t_N) = v_N.$$

In the context of a group action and invariant Lagrangians, we refer the reader to Section 5 for an example of higher-order interpolation particularly relevant for *Computational Anatomy*.

## 2.3 Quotient space and reduced Lagrangian

When one deals with a Lagrangian  $L : T^{(k)}Q \rightarrow \mathbb{R}$  that is invariant with respect to the lift  $\Phi^{(k)} : G \times T^{(k)}Q \rightarrow T^{(k)}Q$  of a group action  $\Phi : G \times Q \rightarrow Q$ , then the invariance can be exploited to define a new function called the *reduced Lagrangian* on the quotient space  $(T^{(k)}Q)/G$ . We review this procedure here. Since this paper mainly deals with the case where  $Q = G$ , we begin by describing this special case.



Let  $G$  be a Lie group and  $h \in G$ . The right-, respectively left-actions by  $h$  on  $G$ ,

$$R_h : G \rightarrow G, \quad g \mapsto gh, \quad \text{and} \quad L_h : G \rightarrow G, \quad g \mapsto hg,$$

can be naturally lifted to actions on the  $k^{\text{th}}$ -order tangent bundle  $T^{(k)}G$  (see (2.2)). We will denote these lifted actions by concatenation, as in

$$\begin{aligned} R_h^{(k)} : T^{(k)}G &\rightarrow T^{(k)}G, & [g]_{g_0}^{(k)} &\mapsto R_h^{(k)}([g]_{g_0}^{(k)}) =: [g]_{g_0}^{(k)}h, & \text{and} \\ L_h^{(k)} : T^{(k)}G &\rightarrow T^{(k)}G, & [g]_{g_0}^{(k)} &\mapsto L_h^{(k)}([g]_{g_0}^{(k)}) =: h[g]_{g_0}^{(k)}. \end{aligned}$$

Consider a Lagrangian  $L : T^{(k)}G \rightarrow \mathbb{R}$  that is right-, or left-invariant, i.e., invariant with respect to the lifted right-, or left-actions of  $G$  on itself. For any  $[g]_{g_0}^{(k)} \in T^{(k)}G$  we then get

$$L([g]_{g_0}^{(k)}) = L|_{T_e^{(k)}G}([g]_{g_0}^{(k)}g_0^{-1}), \quad \text{or} \quad L([g]_{g_0}^{(k)}) = L|_{T_e^{(k)}G}(g_0^{-1}[g]_{g_0}^{(k)}), \quad (2.13)$$

respectively. The restriction  $L|_{T_e^{(k)}G}$  of the Lagrangian to the  $k^{\text{th}}$ -order tangent space at the identity  $e$  therefore fully specifies the Lagrangian  $L$ . Moreover, there are natural identifications  $\alpha_k : T_e^{(k)}G \rightarrow k\mathfrak{g}$  given by

$$\alpha_k([g]_e^{(k)}) := \left( \dot{g}(0), \frac{d}{dt} \Big|_{t=0} \dot{g}(t)g(t)^{-1}, \dots, \frac{d^{k-1}}{dt^{k-1}} \Big|_{t=0} \dot{g}(t)g(t)^{-1} \right), \quad (2.14)$$

or

$$\alpha_k([g]_e^{(k)}) := \left( \dot{g}(0), \frac{d}{dt} \Big|_{t=0} g(t)^{-1}\dot{g}(t), \dots, \frac{d^{k-1}}{dt^{k-1}} \Big|_{t=0} g(t)^{-1}\dot{g}(t) \right), \quad (2.15)$$

respectively, where  $t \mapsto g(t)$  is an arbitrary representative of  $[g]_e^{(k)}$ .

The reduced Lagrangian  $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$  is then defined as

$$\ell := L|_{T_e^{(k)}G} \circ \alpha_k^{-1}, \quad (2.16)$$

where one uses the choice for  $\alpha_k$  that is appropriate, namely (2.14) for a right-invariant Lagrangian  $L$  and (2.15) for a left-invariant Lagrangian  $L$ . Let  $t \mapsto g(t) \in G$  be a curve on the Lie group. For every  $t$  this curve defines an element in  $T_{g(t)}^{(k)}G$ , namely

$$[g]_{g(t)}^{(k)} := [h]_{g(t)}^{(k)}, \quad \text{where } h \text{ is the curve } \tau \mapsto h(\tau) := g(t + \tau). \quad (2.17)$$

Note that for the case  $k = 1$  we write, as usual,  $\dot{g}(t) := [g]_{g(t)}^{(1)}$ . The following lemma is a direct consequence of the definitions:

**Lemma 2.2.** *Let  $t \mapsto g(t)$  be a curve in  $G$  and  $L : T^{(k)}G \rightarrow \mathbb{R}$  a right-, or left-invariant Lagrangian. Then the following equation holds for any time  $t_0$ ,*

$$L([g]_{g(t_0)}^{(k)}) = \ell\left(\xi(t_0), \dot{\xi}(t_0), \dots, \xi^{(k-1)}(t_0)\right), \quad (2.18)$$

where  $\xi := \dot{g}g^{-1}$ , or  $\xi := g^{-1}\dot{g}$  respectively.

This last equation will play a key role in the higher-order Euler-Poincaré reduction discussed in the next section.

### 3 Higher-order Euler-Poincaré reduction

In this section we derive the basic  $k^{\text{th}}$ -order Euler-Poincaré equations by reducing the variational principle associated to the Euler-Lagrange equations on  $T^{(k)}Q$ . The equations adopt a factorized form, in which the Euler-Poincaré operator at  $k = 1$  is applied to the Euler-Lagrange operation acting on the reduced Lagrangian  $\ell(\xi, \dot{\xi}, \ddot{\xi}, \dots, \xi^{(k-1)}) : k\mathfrak{g} \rightarrow \mathbb{R}$  at the given order,  $k$ . We then apply the  $k^{\text{th}}$ -order Euler-Poincaré equations to derive the equations for geometric  $k$ -splines.

#### 3.1 Quotient map, variations and $k^{\text{th}}$ -order Euler-Poincaré equations

Let  $L : T^{(k)}G \rightarrow \mathbb{R}$  be a right-, or left-invariant Lagrangian. Recall from §2.2 that the Euler-Lagrange equations are equivalent to the following variational problem:

For given  $h_i \in G$  and  $[h]_i^{(k-1)} \in T_{h_i}^{(k-1)}G$ ,  $i = 1, 2$ , find a critical curve of the functional

$$\mathcal{J}[g] = \int_{t_1}^{t_2} L\left([g]_{g(t)}^{(k)}\right) dt$$

among all curves  $g : t \in [t_1, t_2] \mapsto g(t) \in G$  satisfying the endpoint condition

$$[g]_{g(t_i)}^{(k-1)} = [h]_i^{(k-1)}, \quad i = 1, 2. \quad (3.1)$$

The time derivatives of up to order  $k - 1$  are therefore fixed at the endpoints, i.e.,  $[g]_{g(t_i)}^{(j)} = [h]_i^{(j)}$ ,  $j = 0, \dots, k - 1$ , are automatically verified. Let  $g : t \mapsto g(t) \in G$  be a curve and  $(\varepsilon, t) \mapsto g_\varepsilon(t) \in G$  a variation of  $g$  respecting (3.1). We recall from Lemma 2.2 that, for any  $\varepsilon$  and any  $t_0$ ,

$$L\left([g_\varepsilon]_{g_\varepsilon(t_0)}^{(k)}\right) = \ell\left(\xi_\varepsilon(t_0), \dots, \xi_\varepsilon^{(k-1)}(t_0)\right), \quad (3.2)$$

where  $\xi_\varepsilon := \dot{g}_\varepsilon g_\varepsilon^{-1}$ , or  $\xi_\varepsilon := g_\varepsilon^{-1} \dot{g}_\varepsilon$  respectively for the right-, or left-invariant Lagrangian  $L$ . The variation  $\delta\xi$  induced by the variation  $\delta g$  is given by

$$\delta\xi = \dot{\eta} \mp [\xi, \eta], \quad (3.3)$$

where  $\eta := (\delta g)g^{-1}$ , or  $\eta := g^{-1}(\delta g)$ , respectively. It follows from the endpoint conditions (3.1) that  $\eta(t_i) = \dot{\eta}(t_i) = \dots = \eta^{(k-1)}(t_i) = 0$  and therefore  $\delta\xi(t_i) = \dots = \partial_t^{k-2}\delta\xi(t_i) = 0$ , for  $i = 1, 2$ .

We are now ready to compute the variation of  $\mathcal{J}$ :

$$\begin{aligned} \delta \int_{t_1}^{t_2} L\left([g]_{g(t)}^{(k)}\right) dt &= \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{t_1}^{t_2} L\left([g_\varepsilon]_{g_\varepsilon(t)}^{(k)}\right) dt \stackrel{(3.2)}{=} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{t_1}^{t_2} \ell\left(\xi_\varepsilon, \dots, \xi_\varepsilon^{(k-1)}\right) dt \\ &= \sum_{j=0}^{k-1} \int_{t_1}^{t_2} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi^{(j)} \right\rangle dt = \sum_{j=0}^{k-1} \int_{t_1}^{t_2} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \partial_t^j \delta \xi \right\rangle dt \\ &= \int_{t_1}^{t_2} \left\langle \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi \right\rangle dt \\ &= \int_{t_1}^{t_2} \left\langle \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \partial_t \eta \mp [\xi, \eta] \right\rangle dt \\ &= \int_{t_1}^{t_2} \left\langle (-\partial_t \mp \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}}, \eta \right\rangle dt, \end{aligned}$$

were we used the vanishing endpoint conditions  $\delta\xi(t_i) = \dots = \partial_t^{k-2}\delta\xi(t_i) = 0$  and  $\eta(t_i) = 0$ , for  $i = 1, 2$ , when integrating by parts. Therefore, the stationarity condition  $\delta\mathcal{J} = 0$  implies the  **$k^{\text{th}}$ -order Euler-Poincaré equation**,

$$(\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta\ell}{\delta\xi^{(j)}} = 0. \quad (3.4)$$

Formula (3.4) takes the following forms for various choices of  $k = 1, 2, 3$ :

If  $k = 1$ :

$$(\partial_t \pm \text{ad}_\xi^*) \frac{\delta\ell}{\delta\xi} = 0,$$

If  $k = 2$ :

$$(\partial_t \pm \text{ad}_\xi^*) \left( \frac{\delta\ell}{\delta\xi} - \partial_t \frac{\delta\ell}{\delta\dot{\xi}} \right) = 0, \quad (3.5)$$

If  $k = 3$ :

$$(\partial_t \pm \text{ad}_\xi^*) \left( \frac{\delta\ell}{\delta\xi} - \partial_t \frac{\delta\ell}{\delta\dot{\xi}} + \partial_t^2 \frac{\delta\ell}{\delta\ddot{\xi}} \right) = 0.$$

The first of these is the usual Euler-Poincaré equation. The others adopt a *factorized* form in which the Euler-Poincaré operator  $(\partial_t \pm \text{ad}_\xi^*)$  is applied to the Euler-Lagrange operation on the reduced Lagrangian  $\ell(\xi, \dot{\xi}, \ddot{\xi}, \dots)$  at the given order.

The results obtained above are summarized in the following theorem.

**Theorem 3.1** ( $k^{\text{th}}$ -order Euler-Poincaré reduction). *Let  $L : T^{(k)}G \rightarrow \mathbb{R}$  be a  $G$ -invariant Lagrangian and let  $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$  be the associated reduced Lagrangian. Let  $g(t)$  be a curve in  $G$  and  $\xi(t) = \dot{g}(t)g(t)^{-1}$ , resp.  $\xi(t) = g(t)^{-1}\dot{g}(t)$  be the reduced curve in the Lie algebra  $\mathfrak{g}$ . Then the following assertions are equivalent.*

- (i) *The curve  $g(t)$  is a solution of the  $k^{\text{th}}$ -order Euler-Lagrange equations for  $L : T^{(k)}G \rightarrow \mathbb{R}$ .*
- (ii) *Hamilton's variational principle*

$$\delta \int_{t_1}^{t_2} L(g, \dot{g}, \dots, g^{(k)}) dt = 0$$

*holds upon using variations  $\delta g$  such that  $\delta g^{(j)}$  vanish at the endpoints for  $j = 0, \dots, k-1$ .*

- (iii) *The  $k^{\text{th}}$ -order Euler-Poincaré equations for  $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ :*

$$(\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta\ell}{\delta\xi^{(j)}} = 0. \quad (3.6)$$

- (iv) *The constrained variational principle*

$$\delta \int_{t_1}^{t_2} \ell(\xi, \dot{\xi}, \dots, \xi^{(k)}) = 0$$

*holds for constrained variations of the form  $\delta\xi = \partial_t\eta \mp [\xi, \eta]$ , where  $\eta$  is an arbitrary curve in  $\mathfrak{g}$  such that  $\eta^{(j)}$  vanish at the endpoints, for all  $j = 0, \dots, k-1$ .*

**Remark 3.2.** The quotient map (2.14), respectively (2.15), can be used for any Lie group  $G$ . In the case of matrix groups, one might consider the alternative quotient map of the form

$$(g, \dot{g}, \dots, g^{(k)}) \rightarrow (\nu_1, \dots, \nu_k), \quad \nu_j := g^{(j)} g^{-1} \quad \text{respectively} \quad \nu_j := g^{-1} g^{(j)}. \quad (3.7)$$

One may easily pass from the variables  $(\xi, \dot{\xi}, \dots, \partial_t^{(k-1)} \xi)$  to the variables  $(\nu_1, \dots, \nu_k)$ . For example:

$$\begin{aligned} \xi &= \nu_1 \\ \dot{\xi} &= \partial_t(\dot{g} g^{-1}) = \ddot{g} g^{-1} - \dot{g} g^{-1} \dot{g} g^{-1} = \nu_2 - \nu_1 \nu_1 \\ \ddot{\xi} &= \nu_3 - 2\nu_2 \nu_1 + 2\nu_1 \nu_1 \nu_1 - \nu_1 \nu_2, \end{aligned} \quad (3.8)$$

and so forth, by using the rule  $\dot{\nu}_j = \nu_{j+1} - \nu_j \nu_1$ . Here all concatenations mean matrix multiplications. One can easily derive the constrained variations and the  $k^{\text{th}}$ -order Euler-Poincaré equations associated to this quotient map in a similar way as above.

### 3.2 Example: Riemannian cubics

In this section we apply the  $k^{\text{th}}$ -order Euler-Poincaré reduction to the particular case of 2-splines on Lie groups. Fix a right-, respectively left-invariant Riemannian metric  $\gamma$  on the Lie group  $G$ . We denote by

$$\|v_g\|_g^2 := \gamma_g(v_g, v_g)$$

the corresponding squared norm of a vector  $v_g \in T_g G$ . The inner product induced on the Lie algebra  $\mathfrak{g}$  is also denoted by  $\gamma : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$  and its squared norm by

$$\|\xi\|_{\mathfrak{g}}^2 := \gamma(\xi, \xi).$$

We recall that the associated isomorphisms

$$\flat : \mathfrak{g} \rightarrow \mathfrak{g}^*, \quad \xi \mapsto \xi^\flat, \quad \text{and} \quad \sharp : \mathfrak{g}^* \rightarrow \mathfrak{g}, \quad \mu \mapsto \mu^\sharp, \quad (3.9)$$

are defined by

$$\langle \xi^\flat, \eta \rangle = \gamma(\xi, \eta), \quad \text{for all } \xi, \eta \in \mathfrak{g}, \quad \text{and} \quad \sharp := \flat^{-1}, \quad (3.10)$$

where  $\langle \cdot, \cdot \rangle$  denotes the dual pairing between  $\mathfrak{g}^*$  and  $\mathfrak{g}$ .

**Proposition 3.3.** *Consider the Lagrangian  $L : T^{(2)}G \rightarrow \mathbb{R}$  for geometric 2-splines, given by*

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2, \quad (3.11)$$

where  $\|\cdot\|$  is the norm of a right-, respectively left-invariant metric on  $G$ . Then  $L$  is right-, respectively left-invariant and induces the reduced Lagrangian  $\ell : 2\mathfrak{g} \rightarrow \mathbb{R}$  given by

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_\xi^\dagger \xi \right\|_{\mathfrak{g}}^2, \quad (3.12)$$

where  $\text{ad}^\dagger$  by  $\text{ad}_\xi^\dagger \eta := (\text{ad}_\xi^*(\eta^\flat))^\sharp$ , for any  $\xi, \eta \in \mathfrak{g}$ .

*Proof.* Let us recall the expression of the Levi-Civita covariant derivative associated to a right (respectively left)  $G$ -invariant Riemannian metric on  $G$ . For  $X \in \mathfrak{X}(G)$  and  $v_g \in T_g G$ , we have (e.g., [KM97], Section 46.5)

$$\nabla_{v_g} X(g) = TR_g \left( \mathbf{d}f(v_g) + \frac{1}{2} \text{ad}_v^\dagger f(g) + \frac{1}{2} \text{ad}_{f(g)}^\dagger v - \frac{1}{2} [v, f(g)] \right), \quad v := v_g g^{-1} \quad (3.13)$$

$$\text{resp. } \nabla_{v_g} X(g) = TL_g \left( \mathbf{d}f(v_g) - \frac{1}{2} \text{ad}_v^\dagger f(g) - \frac{1}{2} \text{ad}_{f(g)}^\dagger v + \frac{1}{2} [v, f(g)] \right), \quad v := g^{-1} v_g \quad (3.14)$$

where  $f \in \mathcal{F}(G; \mathfrak{g})$  is uniquely determined by the condition  $X(g) = TR_g(f(g))$  for right-, respectively  $X(g) = TL_g(f(g))$  for left  $G$ -invariance. Therefore, we have

$$\frac{D}{Dt} \dot{g}(t) = \nabla_{\dot{g}} \dot{g} = TR_g \left( \dot{\xi} + \frac{1}{2} \text{ad}_\xi^\dagger \xi + \frac{1}{2} \text{ad}_\xi^\dagger \xi - \frac{1}{2} [\xi, \xi] \right) = TR_g \left( \dot{\xi} + \text{ad}_\xi^\dagger \xi \right),$$

respectively

$$\frac{D}{Dt} \dot{g}(t) = \nabla_{\dot{g}} \dot{g} = TL_g \left( \dot{\xi} - \frac{1}{2} \text{ad}_\xi^\dagger \xi - \frac{1}{2} \text{ad}_\xi^\dagger \xi + \frac{1}{2} [\xi, \xi] \right) = TL_g \left( \dot{\xi} - \text{ad}_\xi^\dagger \xi \right),$$

where we used  $X(g) = \dot{g}$ ,  $v_g = \dot{g}$ , so  $f(g) = \dot{g}g^{-1} = \xi$  (respectively,  $f(g) = g^{-1}\dot{g} = \xi$ ) and  $\mathbf{d}f(v_g) = \dot{\xi}$ .

Thus we obtain, due to the right-, or left-invariance of the metric  $\gamma$ ,

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2 = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_\xi^\dagger \xi \right\|_{\mathfrak{g}}^2, \quad (3.15)$$

which depends only on the right invariant quantity  $\xi = \dot{g}g^{-1}$ , respectively the left invariant quantity  $\xi = g^{-1}\dot{g}$ . Accordingly,  $L$  is right-, or left-invariant, and the group-reduced Lagrangian is

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \left\| \dot{\xi} \pm \text{ad}_\xi^\dagger \xi \right\|_{\mathfrak{g}}^2$$

which completes the proof.  $\square$

**Remark 3.4.** The above considerations generalize to geometric  $k$ -splines for  $k > 2$ . Indeed, iterated application of formulas (3.13), (3.14) yields

$$\frac{D^k}{Dt^k} \dot{g} = TR_g(\eta_k), \quad \text{respectively} \quad \frac{D^k}{Dt^k} \dot{g} = TL_g(\eta_k),$$

where the quantities  $\eta_k \in \mathfrak{g}$  are defined by the recursive formulae

$$\eta_1 = \dot{\xi} \pm \text{ad}_\xi^\dagger \xi, \quad \text{and} \quad \eta_k = \dot{\eta}_{k-1} \pm \frac{1}{2} \left( \text{ad}_\xi^\dagger \eta_{k-1} + \text{ad}_{\eta_{k-1}}^\dagger \xi + \text{ad}_{\eta_{k-1}} \xi \right), \quad (3.16)$$

for  $\xi = \dot{g}g^{-1}$ , respectively  $\xi = g^{-1}\dot{g}$ . Therefore, the Lagrangian (2.9) for geometric  $k$ -splines on a Lie group  $G$  with right-, respectively left-invariant Riemannian metric,

$$L_k(g, \dot{g}, \dots, g^{(k)}) = \frac{1}{2} \left\| \frac{D^{k-1}}{Dt^{k-1}} \dot{g} \right\|_g^2,$$

is right-, respectively left-invariant, and the reduced Lagrangian is

$$\ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) = \frac{1}{2} \left\| \eta_{k-1} \right\|_{\mathfrak{g}}^2. \quad (3.17)$$

**Computing the second-order Euler-Poincaré equations for splines.** Let us compute the Euler-Poincaré equations for  $k = 2$ . The required variational derivatives of the reduced Lagrangian (3.12) are given by

$$\frac{\delta \ell}{\delta \dot{\xi}} = \dot{\xi}^\flat \pm \text{ad}_\xi^* \xi^\flat =: \eta^\flat \quad \text{and} \quad \frac{\delta \ell}{\delta \xi} = \mp \left( \text{ad}_\eta^* \xi^\flat + (\text{ad}_\eta \xi)^\flat \right) \in \mathfrak{g}^*. \quad (3.18)$$

From formula (3.5) with  $k = 2$  one then finds the  $2^{\text{nd}}$ -order Euler-Poincaré equation

$$(\partial_t \pm \text{ad}_\xi^*) \left( \partial_t \eta^\flat \pm \text{ad}_\eta^* \xi^\flat \pm (\text{ad}_\eta \xi)^\flat \right) = 0, \quad \text{with} \quad \eta^\flat := \dot{\xi}^\flat \pm \text{ad}_\xi^* \xi^\flat, \quad (3.19)$$

or, equivalently,

$$\left( \partial_t \pm \text{ad}_\xi^\dagger \right) \left( \partial_t \eta \pm \text{ad}_\eta^\dagger \xi \pm \text{ad}_\eta \xi \right) = 0, \quad \text{with} \quad \eta := \dot{\xi} \pm \text{ad}_\xi^\dagger \xi. \quad (3.20)$$

These are the reduced equations for geometric 2-splines associated to a left-, or right-invariant Riemannian metric on the Lie group  $G$ .

In an analogous fashion one can derive the Euler-Poincaré equations for geometric  $k$ -splines, using the reduced Lagrangian (3.17).

When the metric is left-, and right-invariant (bi-invariant) further simplifications arise.

**Example 1: Bi-invariant metric and the NHP equation.** In the case of a *bi-invariant* Riemannian metric, we have  $\text{ad}_\xi^\dagger \eta = -\text{ad}_\xi \eta$  and therefore  $\eta^\flat = \dot{\xi}^\flat$ , so that  $\eta = \dot{\xi}$  and the equations (3.20) become

$$(\partial_t \pm \text{ad}_\xi^*) \dot{\xi}^\flat = 0 \quad \text{or} \quad \left( \partial_t \pm \text{ad}_\xi^\dagger \right) \dot{\xi} = 0 \quad \text{or} \quad \ddot{\xi} \mp [\xi, \dot{\xi}] = 0, \quad (3.21)$$

as in [CS95]. Note that in this case, the reduced Lagrangian (3.12) is simply given by  $\ell(\xi, \dot{\xi}) = \frac{1}{2} \|\dot{\xi}\|^2$ . We also remark that since the metric is bi-invariant, one may choose to reduce the system either on the right *or* on the left. This choice will determine which sign appears in (3.21).

Taking  $G = SO(3)$ , we recover the NHP equation (1.1) of [NHP89]:

$$\ddot{\Omega} = \pm \Omega \times \ddot{\Omega}. \quad (3.22)$$

In [NHP89], the unreduced equations in the general case are also derived, but the symmetry reduced equation is given only for  $SO(3)$  with bi-invariant metric.

**Remark 3.5.** [Conventions for  $\mathfrak{so}(3)$  and  $\mathfrak{so}(3)^*$ ]

In equation (3.22) and throughout the paper we use vector notation for the Lie algebra  $\mathfrak{so}(3)$  of the Lie group of rotations  $SO(3)$ , as well as for its dual  $\mathfrak{so}(3)^*$ . One identifies  $\mathfrak{so}(3)$  with  $\mathbb{R}^3$  via the familiar isomorphism

$$\hat{\cdot}: \mathbb{R}^3 \rightarrow \mathfrak{so}(3), \quad \Omega = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \Omega := \hat{\Omega} = \begin{pmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{pmatrix}, \quad (3.23)$$

called the *hat map*. This is a Lie algebra isomorphism when the vector cross product  $\times$  is used as the Lie bracket operation on  $\mathbb{R}^3$ . The identification of  $\mathfrak{so}(3)$  with  $\mathbb{R}^3$  induces an isomorphism of the dual spaces  $\mathfrak{so}(3)^* \cong (\mathbb{R}^3)^* \cong \mathbb{R}^3$ .

**Example 2: Elastica.** Another example of the  $2^{nd}$ -order Euler-Poincaré equation arises in the case of *elastica* treated in [HB04b], whose Lagrangian is

$$L(g, \dot{g}, \ddot{g}) = \frac{\tau^2}{2} \|\dot{g}\|_g^2 + \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \right\|_g^2,$$

and whose reduced Lagrangian is

$$\ell(\xi, \dot{\xi}) = \frac{\tau^2}{2} \|\xi\|_{\mathfrak{g}}^2 + \frac{1}{2} \|\dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi\|_{\mathfrak{g}}^2. \quad (3.24)$$

Using the  $2^{nd}$ -order Euler-Poincaré equation (3.5) one easily obtains the reduced equations

$$\left( \partial_t \pm \text{ad}_{\xi}^{\dagger} \right) \left( \partial_t \eta \pm \text{ad}_{\eta}^{\dagger} \xi \pm \text{ad}_{\eta} \xi - \tau^2 \xi \right) = 0, \quad \text{with} \quad \eta := \dot{\xi} \pm \text{ad}_{\xi}^{\dagger} \xi, \quad (3.25)$$

which simplify to

$$\left( \partial_t \pm \text{ad}_{\xi}^{\dagger} \right) \left( \partial_t^2 \xi - \tau^2 \xi \right) = 0$$

in the bi-invariant case.

**Remark 3.6.** We now consider the particular case  $G = SO(3)$ . Let  $\mathbf{I}$  be a  $3 \times 3$  symmetric positive definite matrix (inertia tensor) and consider the inner product  $\gamma(\Omega_1, \Omega_2) = \mathbf{I}\Omega_1 \cdot \Omega_2$  on  $\mathbb{R}^3$ . The Lagrangian for the elastica on  $SO(3)$  reads

$$L(\Lambda, \dot{\Lambda}, \ddot{\Lambda}) = \frac{\tau^2}{2} \|\dot{\Lambda}\|_{\Lambda}^2 + \frac{1}{2} \left\| \frac{D}{Dt} \dot{\Lambda} \right\|_{\Lambda}^2,$$

where  $\|\cdot\|_{\Lambda}$  is the right-, respectively left-invariant metric induced the inner product  $\gamma$ . Relative to this inner product we have

$$\text{ad}_{\Omega_1}^{\dagger} \Omega_2 = \mathbf{I}^{-1}(\mathbf{I}\Omega_2 \times \Omega_1),$$

so the reduced Lagrangian (3.24) reads

$$\begin{aligned} \ell(\Omega, \dot{\Omega}) &= \frac{\tau^2}{2} \|\Omega\|^2 + \frac{1}{2} \|\dot{\Omega} \pm \mathbf{I}^{-1}(\mathbf{I}\Omega \times \Omega)\|^2 \\ &= \frac{\tau^2}{2} \Omega \cdot \mathbf{I}\Omega + \frac{1}{2} \left( \mathbf{I}\dot{\Omega} \pm \mathbf{I}\Omega \times \Omega \right) \cdot \mathbf{I}^{-1} \left( \mathbf{I}\dot{\Omega} \pm \mathbf{I}\Omega \times \Omega \right). \end{aligned} \quad (3.26)$$

If  $\tau = 0$ , this expression can be interpreted as the Lagrangian for geometric 2-splines of a rigid body.

If  $\mathbf{I}$  is the identity, the Lagrangian in (3.26) simplifies to

$$\ell(\Omega, \dot{\Omega}) = \frac{\tau^2}{2} \Omega \cdot \Omega + \frac{1}{2} \dot{\Omega} \cdot \dot{\Omega}$$

and the Lagrangian of the NHP equation is recovered when  $\tau = 0$ .



**Example 3:  $L^2$ -splines.** One can consider  $L^2$  geometric 2-splines on the diffeomorphism group of a manifold  $\mathcal{D}$  as follows. Fix a Riemannian metric  $g$  on  $\mathcal{D}$  and consider the associated  $L^2$  right-invariant Riemannian metric on  $G = \text{Diff}(\mathcal{D})$  and its induced second-order Lagrangian

$$L(\eta, \dot{\eta}, \ddot{\eta}) = \frac{1}{2} \left\| \frac{D}{Dt} \dot{\eta} \right\|_{\eta}^2$$

on  $T^{(2)} \text{Diff}(\mathcal{D})$ . The reduced Lagrangian on  $2\mathfrak{g} = 2\mathfrak{X}(\mathcal{D})$  reads

$$\ell(\mathbf{u}, \dot{\mathbf{u}}) = \frac{1}{2} \|\dot{\mathbf{u}} + \text{ad}_{\mathbf{u}}^{\dagger} \mathbf{u}\|^2,$$

where  $\text{ad}^{\dagger}$  denotes the transpose with respect to the  $L^2$  inner product, given by  $\text{ad}_{\mathbf{u}}^{\dagger} \mathbf{v} = \nabla_{\mathbf{u}} \mathbf{v} + (\nabla \mathbf{u})^{\top} \cdot \mathbf{v} + \mathbf{v} \text{div } \mathbf{u}$ . In this case, the  $\text{ad}^{\dagger}$  and  $\text{ad}$  terms in (3.20) combine to produce the spline equation

$$(\partial_t + \text{ad}_{\mathbf{u}}^{\dagger}) (\partial_t \mathbf{v} + 2\mathbf{S} \mathbf{v} \cdot \mathbf{u} + \mathbf{u} \text{div } \mathbf{v}) = 0, \quad \mathbf{v} = \partial_t \mathbf{u} + 2\mathbf{S} \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \text{div } \mathbf{u},$$

where  $\mathbf{S} \mathbf{u} := (\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}) / 2$  is the strain-rate tensor.

In the incompressible case, that is when  $G = \text{Diff}_{\text{vol}}(\mathcal{D})$ , the transpose of  $\text{ad}$  relative to the  $L^2$  inner product on divergence free vector field is denoted by  $\text{ad}^+$  and is related to  $\text{ad}^{\dagger}$  by the formula

$$\text{ad}_{\mathbf{u}}^+ \mathbf{v} = \mathbb{P} (\text{ad}_{\mathbf{u}}^{\dagger} \mathbf{v}) = \mathbb{P} (\nabla_{\mathbf{u}} \mathbf{v} + (\nabla \mathbf{u})^{\top} \cdot \mathbf{v}),$$

where  $\mathbb{P}$  denotes the Hodge projector onto the divergence free vector fields. In this case (3.20) reads

$$(\partial_t + \text{ad}_{\mathbf{u}}^+) (\partial_t \mathbf{v} + \mathbb{P} (\nabla_{\mathbf{v}} \mathbf{u} + (\nabla \mathbf{v})^{\top} \cdot \mathbf{u}) + \nabla_{\mathbf{u}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{u}) = 0, \quad \mathbf{v} = \partial_t \mathbf{u} + 2\mathbb{P} (\mathbf{S} \mathbf{u} \cdot \mathbf{u}), \quad \text{div } \mathbf{u} = 0.$$

Remarkably, using the formula  $\text{ad}_{\mathbf{u}}^{\dagger} \nabla p = \nabla (\nabla p \cdot \mathbf{u})$  for  $\text{div } \mathbf{u} = 0$ , all the gradient terms arising from the Hodge projector can be assembled in a single gradient term, thereby producing the incompressible 2-spline equations

$$(\partial_t + \text{ad}_{\mathbf{u}}^{\dagger}) (\partial_t \mathbf{w} + 2\mathbf{S} \mathbf{w} \cdot \mathbf{u}) = -\nabla p, \quad \mathbf{w} = \partial_t \mathbf{u} + 2\mathbf{S} \mathbf{u} \cdot \mathbf{u}, \quad \text{div } \mathbf{u} = 0,$$

where  $\text{ad}^{\dagger}$  (and not  $\text{ad}^+$ ) is used.

**Example 4:  $H^1$ -splines.** One can alternatively consider splines relative to the right-invariant metric induced by the  $H^1$  inner product  $\langle Q \mathbf{u}, \mathbf{u} \rangle$ , where  $Q = (1 - \alpha^2 \Delta)$ . In this case, the 2-spline equation reads

$$(\partial_t + \text{ad}_{\mathbf{u}}^{\dagger}) (\partial_t Q \mathbf{v} + 2\mathbf{S}(Q \cdot \text{ad}_{\mathbf{v}}) \cdot \mathbf{u}) = 0, \quad Q \mathbf{v} = \partial_t Q \mathbf{u} + \text{ad}_{\mathbf{u}}^{\dagger} Q \mathbf{u} \quad (3.27)$$

where  $\mathbf{S}(L) := \frac{1}{2}(L + L^*)$ .

**Remark 3.7.** Note that in order to obtain the simple expression

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \|\xi\|_1^2 + \frac{1}{2} \|\dot{\xi}\|_2^2$$

(instead of (3.24)) for  $\ell$  by reduction, one needs to modify the spline Lagrangian as

$$L(g, \dot{g}, \ddot{g}) = \frac{1}{2} \|\dot{g}\|_1^2 + \frac{1}{2} \left\| \frac{D}{Dt} \dot{g} \pm \text{ad}_g^\dagger \dot{g} \right\|_2^2,$$

where  $\|\cdot\|_i$ ,  $i = 1, 2$ , are two norms associated to two  $G$ -invariant Riemannian metrics on  $G$  and  $\text{ad}^\dagger$  is extended as a bilinear map  $T_g G \times T_g G \rightarrow T_g G$  by  $G$ -invariance.

The associated  $2^{nd}$  order Euler-Poincaré equations are simpler than the one associated to 2-splines. For example, the reduced Lagrangian  $\ell(\mathbf{u}, \dot{\mathbf{u}}) = \frac{1}{2} \langle (1 - \alpha^2 \Delta) \mathbf{u}, \mathbf{u} \rangle + \frac{1}{2} \|\dot{\mathbf{u}}\|^2$  produces the following modification of the EPDiff equation:

$$(\partial_t + \text{ad}_\mathbf{u}^\dagger) (\mathbf{u} - (\alpha^2 \Delta \mathbf{u} + \mathbf{u}_{tt})) = 0. \quad (3.28)$$

### 3.3 Parameter dependent Lagrangians

In many situations, such as the heavy top or the compressible fluid, the Lagrangian of the system is defined on the tangent bundle  $TG$  of the configuration Lie group  $G$ , but it is not  $G$ -invariant. In these cases, the Lagrangian depends parametrically on a quantity  $q_0$  in a manifold  $Q$  on which  $G$  acts and that breaks the symmetry of the Lagrangian  $L = L_{q_0}$ . We refer to [HMR98], [GBR09] for the case of (affine) representation on vector spaces, relation with semidirect products and many examples. This theory was extended to arbitrary actions on manifolds in [GBT10] for applications to symmetry breaking phenomena. We now briefly present the extension of this theory to the case of higher order Lagrangians.

Consider a  $k^{th}$ -order Lagrangian  $L_{q_0} : T^{(k)}G \rightarrow \mathbb{R}$ , depending on a parameter  $q_0$  in a manifold  $Q$ . We suppose that  $G$  acts on the manifold  $Q$  and that the Lagrangian  $L$  is  $G$ -invariant under the action of  $G$  on both  $T^{(k)}G$  and  $Q$ , where we now see  $L$  as a function defined on  $T^{(k)}G \times Q$ . Concerning the action of  $G$  on  $T^{(k)}G \times Q$ , there are several variants that one needs to consider, since they all appear in applications.

(1) First, one has the right, respectively the left action

$$\begin{aligned} (g, \dot{g}, \dots, g^{(k)}, q_0) &\mapsto (gh, \dot{gh}, \dots, g^{(k)}h, h^{-1}q_0), \\ \text{respectively } (g, \dot{g}, \dots, g^{(k)}, q_0) &\mapsto (hg, h\dot{g}, \dots, hg^{(k)}, q_0h^{-1}). \end{aligned}$$

The reduced variables are  $(\xi, q) = (\dot{g}g^{-1}, gq_0)$ , respectively  $(\xi, q) = (g^{-1}\dot{g}, q_0g)$ . In this case, the  $k^{th}$ -order Euler-Lagrange equations for  $L_{q_0}$  on  $T^{(k)}G$  (where  $q_0 \in Q$  is a fixed parameter) are equivalent to the  $k^{th}$ -order Euler-Poincaré equations together with the advection equation

$$(\partial_t \pm \text{ad}_\xi^*) \left( \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \right) = \mathbf{J} \left( \frac{\delta \ell}{\delta q} \right), \quad \partial_t q - \xi_Q(q) = 0, \quad (3.29)$$

with initial condition  $q_0$ . Here  $\xi_Q(q)$  denotes the infinitesimal generator of the  $G$  action on  $Q$  and  $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$  defined by  $\langle \mathbf{J}(\alpha_q), \xi \rangle := \langle \alpha_q, \xi_Q(q) \rangle$  denotes the momentum map associated to the  $G$  action on  $T^*Q$ .

The associated variational principle reads

$$\delta \int_{t_1}^{t_2} \ell \left( \xi, \dot{\xi}, \dots, \xi^{(k-1)}, q \right) dt = 0,$$

relative to the constrained variations (3.3) and constrained variations of  $q$  given by  $\delta q = \eta_Q(q)$ , where  $\eta = (\delta g)g^{-1}$ , respectively  $\eta = g^{-1}(\delta g)$ . Equations (3.29) and their variational formulation can be obtained by an easy generalization of the approach used in Section §3.1.

(2) Secondly, one can consider the right, respectively the left action

$$\begin{aligned} (g, \dot{g}, \dots, g^{(k)}, q_0) &\mapsto (gh, \dot{g}h, \dots, g^{(k)}h, q_0h), \\ \text{respectively } (g, \dot{g}, \dots, g^{(k)}, q_0) &\mapsto (hg, h\dot{g}, \dots, hg^{(k)}, hq_0). \end{aligned}$$

The reduced variables are  $(\xi, q) = (\dot{g}g^{-1}, q_0g^{-1})$ , respectively  $(\xi, q) = (g^{-1}\dot{g}, g^{-1}q_0)$  and one gets the reduced equations

$$(\partial_t \pm \text{ad}_\xi^*) \left( \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \right) = -\mathbf{J} \left( \frac{\delta \ell}{\delta q} \right), \quad \partial_t q + \xi_Q(q) = 0, \quad (3.30)$$

with initial condition  $q_0$ .

If  $Q = V^*$  is the dual of a  $G$ -representation space so which  $G$  acts on  $V^*$  by the dual representation, the above equations reduce to

$$(\partial_t \pm \text{ad}_\xi^*) \left( \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \right) = \frac{\delta \ell}{\delta a} \diamond a, \quad (3.31)$$

where the diamond operation  $\diamond : V \times V^* \rightarrow \mathfrak{g}^*$  is defined by

$$\langle v \diamond a, \xi \rangle = \langle a, \xi_V(v) \rangle,$$

and therefore  $\mathbf{J}(a, v) = -v \diamond a$ . These are the higher order version of the Euler-Poincaré equations with advected quantities studied in [HMR98].

**Example : Rate-dependent fluid models.** Rate-dependent fluid models are usually defined using Lagrangians that depend on the strain-rate tensor  $\mathbf{S} := (\nabla \mathbf{u} + (\nabla \mathbf{u})^\top)/2$  and its higher *spatial* derivatives [BFHL88]. A related class of spatially-regularized fluid models have been introduced as turbulence models [FHT01].

Yet another class of rate-dependent fluid models may be defined, e.g., as *2nd* order Euler-Lagrange equations  $T^{(2)} \text{Diff}(\mathcal{D})$  for a parameter dependent Lagrangian  $L_{a_0}$ . The group reduced representation of the equations of motion for such rate-dependent fluids is found from the previous manipulations, namely

$$(\partial_t \pm \mathcal{L}_{\mathbf{u}}) \left( \frac{\delta \ell}{\delta \mathbf{u}} - \partial_t \frac{\delta \ell}{\delta \dot{\mathbf{u}}} \right) = \frac{\delta \ell}{\delta a} \diamond a, \quad (\partial_t + \mathcal{L}_{\mathbf{u}}) a = 0. \quad (3.32)$$

One has the  $+$  sign for right and the  $-$  sign for left invariance. The usual Eulerian fluid representation is right-invariant and so takes the  $+$  sign. Physically, these fluid models penalize the flow for producing higher temporal frequencies. Therefore, these models might be considered as candidates for *frequency-regularized* models for fluid turbulence. The Kelvin theorem for these fluids involves circulation of the higher time derivatives. For right-invariant higher-order Lagrangians, the Kelvin theorem becomes

$$\frac{d}{dt} \oint_{c(\mathbf{u})} \frac{1}{D} \left( \frac{\delta \ell}{\delta \mathbf{u}} - \partial_t \frac{\delta \ell}{\delta \dot{\mathbf{u}}} \right) = \oint_{c(\mathbf{u})} \frac{1}{D} \frac{\delta \ell}{\delta a} \diamond a,$$

where the density  $D$  satisfies the continuity equation  $(\partial_t + \mathcal{L}_{\mathbf{u}})D = 0$ . Consequently, the integrands in the previous formula are 1-forms and thus may be integrated around the closed curve  $c(\mathbf{u})$  moving with the fluid velocity,  $\mathbf{u}$ . This is the statement of the Kelvin-Noether theorem [HMR98] for  $k$ -splines.

### 3.4 Splines with constraints

Suppose that one wants to minimize the action  $\int_{t_0}^{t_1} L(q, \dot{q}, \ddot{q}) dt$  over curves  $q(t) \in Q$  subject to the condition

$$\langle \omega_i(q), \dot{q} \rangle = k_i, \quad i = 1, \dots, k,$$

where  $\omega_i$  are 1-forms and  $k_i \in \mathbb{R}$ . One uses the variational principle

$$\delta \int_{t_0}^{t_1} \left( L(q, \dot{q}, \ddot{q}) + \sum_{i=1}^k \lambda_i (\langle \omega_i(q), \dot{q} \rangle - k_i) \right) dt = 0, \quad (3.33)$$

for arbitrary variations  $\delta \lambda_i$  of the curves  $\lambda_i(t)$ ,  $i = 1, \dots, k$ , and for variations  $\delta q$  vanishing at the endpoints. Variations relative to  $q$  yield the equation

$$\frac{d^2}{dt^2} \frac{\partial L}{\partial \ddot{q}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q} = \sum_{i=1}^k \left( \lambda_i \mathbf{i}_{\dot{q}} d\omega_i + \dot{\lambda}_i \omega_i \right),$$

whereas variations relative to  $\lambda_i$  yield the constraint. For example, for the Lagrangian (2.8) this yields the equations

$$\frac{D^3}{Dt^3} \dot{q}(t) + R \left( \frac{D}{Dt} \dot{q}(t), \dot{q}(t) \right) \dot{q}(t) = \tau^2 \frac{D}{Dt} \dot{q}(t) + \sum_{i=1}^k \left( \lambda_i \mathbf{i}_{\dot{q}} d\omega_i + \dot{\lambda}_i \omega_i \right),$$

as in [BC93], [CS95], [HB04a], where  $\omega_i = X_i^\flat$  for given linearly independent vector fields  $X_1, \dots, X_k \in \mathfrak{X}(Q)$ .

**Remark 3.8.** The variational principle (3.33) is equivalent to

$$\delta \int_{t_0}^{t_1} \left( L(q, \dot{q}, \ddot{q}) + \sum_{i=1}^k \lambda_i \langle \omega_i(q), \dot{q} \rangle \right) dt = 0 \quad \text{and} \quad \langle \omega_i(q), \dot{q} \rangle = k_i, \quad i = 1, \dots, k, \quad (3.34)$$

where only variations of  $q$  are involved and the term containing  $k_i$  is suppressed.

We now consider the special case  $Q = G$  and we suppose that the one-forms  $\omega_i$ ,  $i = 1, \dots, k$ , are  $G$ -invariant. That is, we can write

$$\langle \omega_i(g), v_g \rangle = \langle \omega_i(g)g^{-1}, v_g g^{-1} \rangle = \langle \zeta_i, v_g g^{-1} \rangle \quad \text{resp.} \quad \langle \omega_i(g), v_g \rangle = \langle \zeta_i, g^{-1}v_g \rangle,$$

where

$$\zeta_i := \omega_i(e) \in \mathfrak{g}^*.$$

The reduction of the variational principle (3.33) yields the constrained variational principle

$$\delta \int_{t_0}^{t_1} \left( \ell(\xi, \dot{\xi}) + \sum_{i=1}^k \lambda_i (\langle \zeta_i, \xi \rangle - k_i) \right) dt = 0,$$

for arbitrary variations  $\delta \lambda_i$  of the curves  $\lambda_i(t)$ ,  $i = 1, \dots, k$ , and variations of  $\xi(t)$  satisfying the constraints (3.3). Equivalently, using (3.34) we rewrite the stationarity condition as

$$\delta \int_{t_0}^{t_1} \left( \ell(\xi, \dot{\xi}) + \langle z, \xi \rangle \right) dt = 0 \quad \text{and} \quad \langle \zeta_i, \xi \rangle = k_i, \quad i = 1, \dots, k,$$

for variations of  $\xi$  satisfying (3.3) and where we have defined  $z := \sum_{i=1}^k \lambda_i \zeta_i \in \mathfrak{g}^*$ . We obtain the equations

$$(\partial_t \pm \text{ad}_\xi^*) \left( \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} + z \right) = 0, \quad \langle \zeta_i, \xi \rangle = k_i.$$

With the Lagrangian (3.12) for 2-splines, we find the reduced equations

$$\left( \partial_t \pm \text{ad}_\xi^\dagger \right) \left( \partial_t \eta \pm \text{ad}_\eta^\dagger \xi \pm \text{ad}_\eta \xi - z \right) = 0, \quad \text{with} \quad \eta := \dot{\xi} \pm \text{ad}_\xi^\dagger \xi$$

and for bi-invariant metrics, we get

$$\left( \partial_t \pm \text{ad}_\xi^\dagger \right) \left( \ddot{\xi} - z \right) = 0, \quad \text{i.e.,} \quad \ddot{\xi} \mp [\xi, \ddot{\xi} - z] - \dot{z} = 0, \quad \langle \zeta_i, \xi \rangle = k_i,$$

which coincides with equation (39) in [CS95]. See also [BC96a].

We can also consider higher-order constraints, with associated variational principle

$$\delta \int_{t_0}^{t_1} \left( \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) + \langle z_0, \xi \rangle + \dots + \langle z_{k-1}, \xi^{(k-1)} \rangle \right) dt = 0.$$

In this case, one obtains the equations

$$(\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \left( \frac{\delta \ell}{\delta \xi^{(j)}} + z_j \right) = 0.$$

For example, with  $k = 2$ , and a bi-invariant metric we have

$$\left( \partial_t \pm \text{ad}_\xi^\dagger \right) \left( \ddot{\xi} + \dot{z}_1 - z_0 \right) = 0.$$

## 4 Clebsch-Pontryagin optimal control

Here we develop the  $k^{\text{th}}$ -order Euler-Poincaré equations from an optimal control approach. The ideas in [GBR10] for  $k = 1$  generalize easily to higher order.

**Definition 4.1.** *Let  $\Phi$  be a (right or left) action of a Lie group  $G$  on a manifold  $Q$ . For a Lie algebra element  $\xi \in \mathfrak{g}$  let*

$$\xi_Q(q) := \left. \frac{d}{dt} \right|_{t=0} \Phi_{\exp(t\xi)}(q),$$

*denote the corresponding infinitesimal generator of the action. Given a cost function  $\ell : k\mathfrak{g} \rightarrow \mathbb{R}$ , the **Clebsch-Pontryagin optimal control problem** is, by definition,*

$$\min_{\xi(t)} \int_0^T \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) dt \quad (4.1)$$

*subject to the following conditions:*

- (A)  $\dot{q} = \xi_Q(q)$  or (A)'  $\dot{q} = -\xi_Q(q)$ ;
- (B)  $q(0) = q_0$  and  $q(T) = q_T$ ;
- (C)  $\xi^{(j)}(0) = \xi_0^j$  and  $\xi^{(j)}(T) = \xi_T^j$ ,  $j = 0, \dots, k-2$ ,

*where  $q_0, q_T \in Q$  and  $\xi_0^j, \xi_T^j \in \mathfrak{g}$ ,  $j = 0, \dots, k-2$ , are given.*

**Variational equations.** We suppose that condition (A) of Definition 4.1 holds. (The calculation for case (A)' is similar.) The resolution of this problem uses the Pontryagin maximum principle which, under sufficient smoothness conditions, implies that its solution necessarily satisfies the variational principle

$$\delta \int_0^T \left( \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}) + \langle \alpha, \dot{q} - \xi_Q(q) \rangle \right) dt = 0,$$

for curves  $t \mapsto \xi(t) \in \mathfrak{g}$  and  $t \mapsto \alpha(t) \in T_{q(t)}^*Q$ . This variational principle yields the conditions

$$\mathbf{J}(\alpha(t)) = \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \quad \text{and} \quad \dot{\alpha} = \xi_{T^*Q}(\alpha), \quad (4.2)$$

in which  $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$  is the cotangent bundle momentum map, as above, and  $\xi_{T^*Q}$  denotes the infinitesimal generator of the cotangent lifted action, denoted  $\Phi^{T^*} : G \times T^*Q \rightarrow T^*Q$ .

If  $G$  acts on the right (respectively left), a solution  $\alpha(t)$  of  $\dot{\alpha} = \xi_{T^*Q}(\alpha)$  is necessarily of the form  $\alpha(t) = \Phi_{g(t)}^{T^*}(\alpha(0))$ , where  $g(0) = e$  and  $\dot{g}(t)g(t)^{-1} = \xi(t)$  (respectively  $g(t)^{-1}\dot{g}(t) = \xi(t)$ ). The above conditions imply coadjoint motion,

$$\mathbf{J}(\alpha(t)) = \mathbf{J}(\Phi_{g(t)}^{T^*}(\alpha(0))) = \text{Ad}_{g(t)}^* \mathbf{J}(\alpha(0)), \quad \text{respectively} \quad \mathbf{J}(\alpha(t)) = \text{Ad}_{g(t)^{-1}}^* \mathbf{J}(\alpha(0)),$$

and by differentiating relative to time, we obtain the left (right) Euler-Poincaré equations:

$$\frac{d}{dt} \mathbf{J}(\alpha(t)) = \text{ad}_{g(t)^{-1}\dot{g}(t)}^* \mathbf{J}(\alpha(t)) = \text{ad}_{\xi(t)}^* \mathbf{J}(\alpha(t)),$$

respectively

$$\frac{d}{dt} \mathbf{J}(\alpha(t)) = -\operatorname{ad}_{\dot{g}(t)g(t)^{-1}}^* \mathbf{J}(\alpha(t)) = -\operatorname{ad}_{\xi(t)}^* \mathbf{J}(\alpha(t)).$$

Upon using the first condition in (4.2), we recover the  $k^{\text{th}}$ -order Euler-Poincaré equations,

$$(\partial_t \mp \operatorname{ad}_{\xi}^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = 0. \quad (4.3)$$

**Example 1: Clebsch approach to the NHP equations.** The NHP equations can be obtained from the Clebsch approach by considering the action of  $SO(3)$  on  $\mathbb{R}^3$ . The Clebsch-Pontryagin control problem is

$$\min_{\xi(t)} \int_0^T \|\dot{\mathbf{Q}}\|^2 dt, \quad \text{subject to} \quad \dot{\mathbf{q}} = \mathbf{\Omega} \times \mathbf{q}, \quad \mathbf{q}(0) = \mathbf{q}_0, \quad \mathbf{q}(T) = \mathbf{q}_T, \quad \mathbf{\Omega}(0) = \mathbf{\Omega}_0, \quad \mathbf{\Omega}(T) = \mathbf{\Omega}_T.$$

The stationarity conditions (4.2) read  $\mathbf{q} \times \mathbf{p} = -\ddot{\mathbf{Q}}$ ,  $\dot{\mathbf{q}} = \mathbf{\Omega} \times \mathbf{q}$  and  $\dot{\mathbf{p}} = \mathbf{\Omega} \times \mathbf{p}$ . One directly observes that they imply the NHP equations.

**Example 2: Clebsch approach to  $H^1$ -splines.** We let the diffeomorphism group  $\operatorname{Diff}_{\text{vol}}(\mathcal{D})$  act on the left on the space of embeddings  $\operatorname{Emb}(S, \mathcal{D})$  of a manifold  $S$  in  $D$ . The associated Clebsch-Pontryagin control problem is

$$\min_{\mathbf{u}(t)} \int_0^T \|\dot{\mathbf{u}} + \operatorname{ad}_{\mathbf{u}}^\dagger \mathbf{u}\|_{H^1}^2 dt, \quad \text{subject to} \quad \dot{\mathbf{Q}} = \mathbf{u} \circ \mathbf{Q}, \quad \mathbf{Q}(0) = \mathbf{Q}_0, \quad \mathbf{Q}(T) = \mathbf{Q}_T, \\ \mathbf{u}(0) = \mathbf{u}_0, \quad \mathbf{u}(T) = \mathbf{u}_T.$$

The condition

$$\mathbf{J}(\mathbf{Q}, \mathbf{P}) = \partial_t Q \mathbf{v} + 2S(Q \cdot \operatorname{ad}_{\mathbf{v}}) \cdot \mathbf{u},$$

where  $Q \mathbf{v} = \partial_t Q \mathbf{u} + \operatorname{ad}_{\mathbf{u}}^\dagger Q \mathbf{u}$ , together with the Hamilton equations on  $T^* \operatorname{Emb}(S, \mathcal{D})$  imply the  $H^1$  spline equations (3.27). In the case of equations (3.28) the condition (4.2) reads

$$\mathbf{J}(\mathbf{Q}, \mathbf{P}) = Q \mathbf{u} - \mathbf{u}_{tt}.$$

**Additional  $q$ -dependence in the Lagrangian.** One can easily include a  $q$ -dependence in the cost function  $\ell$  of the Clebsch-Pontryagin optimal control problem (4.1). In this case, the stationarity conditions (4.2) become

$$\mathbf{J}(\alpha(t)) = \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} \quad \text{and} \quad \dot{\alpha} = \xi_{T^*Q}(\alpha) + \operatorname{Ver}_{\alpha} \frac{\delta \ell}{\delta q}, \quad (4.4)$$

where for  $\alpha, \beta \in T_q^*Q$ , the *vertical lift of  $\beta$  relative to  $\alpha$*  is defined by

$$\operatorname{Ver}_{\alpha} \beta := \left. \frac{d}{ds} \right|_{s=0} (\alpha + s\beta) \in T_{\alpha}(T^*Q).$$



In this case, the differential equation (4.3) for the control  $\xi(t)$  generalizes to

$$(\partial_t \mp \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = \mathbf{J} \left( \frac{\delta \ell}{\delta q} \right), \quad (4.5)$$

where  $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$  is again the cotangent bundle momentum map.

**Remark 4.2.** [Recovering Euler-Poincaré equations of §3.3]

Equations (4.5) for the control recover the  $k^{\text{th}}$ -order Euler-Poincaré equations (3.29). Note that a right, respectively left action of  $G$  on  $Q$  produces the left, respectively right Euler-Poincaré equations in (4.3) consistently with the results in §3.3. In order to obtain the Euler-Poincaré equations (3.30) one needs to impose condition (A)' instead of (A) on the dynamics on  $q$ .

## 5 Higher-order template matching problems

In this section we generalize the methods of [BGBHR10] to higher order because the added smoothness provided by higher-order models makes them attractive for longitudinal data interpolation, in particular in Computational Anatomy (CA).

We first give a brief account of the previous work done on longitudinal data interpolation in CA. Then we derive the equations that generalize [BGBHR10]. After making a few remarks concerning the gain in smoothness, we provide a qualitative discussion of two Lagrangians of interest for CA. Finally, we close the section by demonstrating the spline approach to template matching for the finite dimensional case of fitting a spline through a sequence of orientations on  $SO(3)$ .

### 5.1 Previous work on longitudinal data interpolation in CA

CA is concerned with modeling and quantifying diffeomorphic evolutions of shapes, as presented in [MTY02, MY01]. Usually one aims at finding a geodesic path, on the space of shapes, between given initial and final data. This approach can be adapted for longitudinal data interpolation; that is, interpolation through a sequence of data points. One may interpolate between the given data points in such a way that the path is piecewise-geodesic, [BK08, DPG<sup>+</sup>09]. It was, however, argued in [TV10] that higher order models, i.e., models that provide more smoothness than the piecewise-geodesic one, are better suited as growth models for typical biological evolutions. As an example of such a higher-order model, spline interpolation on the Riemannian manifold of landmarks was studied there. In the next paragraph, we will consider another class of models of interest for CA that are inspired by an optimal control viewpoint. Indeed, the time-dependent vector field can be seen as a control variable acting on the template and the penalization on this control variable will be directly defined on the Lie algebra. Finally, we underline that this class of models is an interesting alternative to the shape splines model presented in [TV10].

### 5.2 Euler-Lagrange equations for higher-order template matching

Let  $G$  be a Lie group with Lie algebra  $\mathfrak{g}$ , and let

$$G \times V \rightarrow V, \quad (g, I) \mapsto gI \quad (5.1)$$

be a left representation of  $G$  on  $V$ . Let  $\|\cdot\|_V$  be a norm on  $V$ . We consider minimization problems of the following abstract form:

Given a Lagrangian  $\ell : (k-1)\mathfrak{g} \rightarrow \mathbb{R}$ ,  $\sigma, t_1, \dots, t_l \in \mathbb{R}$ ,  $T_0, I_{t_1}, \dots, I_{t_l} \in V$ , and  $\xi_0^0, \dots, \xi_0^{k-2} \in \mathfrak{g}$ , minimize the functional

$$E[\xi] := \int_0^{t_l} \ell(\xi(t), \dots, \xi^{(k-1)}(t)) dt + \frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \quad (5.2)$$

subject to the conditions  $\xi^{(j)}(0) = \xi_0^j$ ,  $j = 0, \dots, k-2$ , where  $g^\xi(t_i)$  is the flow of  $\xi(t)$  evaluated at time  $t_i$ . The minimization is carried out over the space

$$\mathcal{P}_{k-1} := \{\xi \in C^{k-2}([0, t_l], \mathfrak{g}) \mid \xi^{(k-1)} \in C_{pcw}^\infty([t_i, t_{i+1}])_{i=0}^{l-1}\}$$

where  $C_{pcw}^\infty([t_i, t_{i+1}])_{i=0}^{l-1}$  denotes the set of piecewise smooth curves whose only discontinuities would be at the  $t_i$ ,  $i = 1, \dots, l-1$ , i.e.

$$C_{pcw}^\infty([t_i, t_{i+1}])_{i=0}^{l-1} := \left\{ f \in L^2([0, T], \mathfrak{g}) \mid \forall i = 0, \dots, l-1 \exists f^i \in C^\infty([t_i, t_{i+1}], \mathfrak{g}) \text{ s.t. } f = f^i_{|[t_i, t_{i+1}]}\right\}.$$

More precisely, given such a curve  $\xi(t)$  in the Lie algebra  $\mathfrak{g}$ , its flow  $g^\xi : t \mapsto g^\xi(t) \in G$  is a continuous curve defined by the conditions

$$g^\xi(0) = e, \quad \text{and} \quad \frac{d}{dt}g^\xi(t) = \xi(t)g^\xi(t), \quad (5.3)$$

whenever  $t$  is in one of the open intervals  $(0, t_1), \dots, (t_{l-1}, t_l)$ . Here we used the notation  $\xi(t)g^\xi(t) := TR_{g^\xi(t)}\xi(t)$ .

We typically think of  $(I_{t_1}, \dots, I_{t_l})$  as the time-sequence of data, indexed by time points  $t_j$ ,  $j = 1, \dots, l$ , and  $T_0$  is the template (the source image). Moreover,  $\xi : t \in [0, t_l] \mapsto \xi(t) \in \mathfrak{g}$  is typically a time-dependent vector field (sufficiently smooth in time) that generates a flow of diffeomorphisms  $g^\xi : t \in [0, t_l] \mapsto g^\xi(t) \in G$ . Note that, in this case, the Lie group  $G$  is infinite dimensional and a rigorous framework to work in is the *large deformations by diffeomorphisms* setting thoroughly explained in [TY05]. We will informally refer to this case as the *diffeomorphisms case* or *infinite dimensional case*. The expression  $g^\xi(t_i)T_0$  represents the template at time  $t_i$ , as it is being deformed by the flow of diffeomorphisms. Inspired by the second-order model presented in [TV10], this subsection thus generalizes the work of [BGBHR10] in two directions. First, we allow for a *higher-order* penalization on the time-dependent vector field given by the first term of the functional (5.2); second, the similarity measure (second term in (5.2)) takes into account *several* time points in order to compare the deformed template with the time-sequence target.

Staying at a general level, we will take the geometric viewpoint of [BGBHR10] in order to derive the Euler-Lagrange equations satisfied by any minimizer of  $E$ . We suppose that the norm on  $V$  is induced by an inner product  $\langle \cdot, \cdot \rangle_V$  and denote by  $\flat$  the isomorphism

$$\flat : V \rightarrow V^*, \quad \omega \mapsto \omega^\flat$$

that satisfies

$$\langle I, J \rangle_V = \langle I^\flat, J \rangle \quad \text{for all } I, J \in V,$$

where we wrote  $\langle \cdot, \cdot \rangle$  for the duality pairing between  $V$  and its dual  $V^*$ . The action (5.1) of  $G$  on  $V$  induces an action on  $V^*$ ,

$$G \times V^* \rightarrow V^*, \quad (g, \omega) \mapsto g\omega$$

that is defined by the identity

$$\langle g\omega, I \rangle = \langle \omega, g^{-1}I \rangle \quad \text{for all } I \in V, \omega \in V^*, g \in G. \quad (5.4)$$

The cotangent-lift momentum map  $\diamond : V \times V^* \rightarrow \mathfrak{g}^*$  for the action of  $G$  on  $V$  is defined by the identity

$$\langle I \diamond \omega, \xi \rangle = \langle \omega, \xi I \rangle, \quad \text{for all } I \in V, \omega \in V^*, \xi \in \mathfrak{g}, \quad (5.5)$$

where the brackets on both sides represent the duality pairings of the respective spaces for  $\mathfrak{g}$  and  $V$ , and where  $\xi I$  denotes the infinitesimal action of  $\mathfrak{g}$  on  $V$ , defined as  $\xi I := \left. \frac{d}{dt} \right|_{t=0} g(t)I \in V$  for any  $C^1$  curve  $g : [-\varepsilon, \varepsilon] \rightarrow G$  that satisfies  $g(0) = e$  and  $\left. \frac{d}{dt} \right|_{t=0} g(t) = \xi \in \mathfrak{g}$ . Note that equations (5.4) and (5.5) imply

$$\text{Ad}_{g^{-1}}^*(I \diamond \omega) = gI \diamond g\omega. \quad (5.6)$$

For the flow defined in (5.3), we also introduce the notation

$$g_{t,s}^\xi := g^\xi(t) (g^\xi(s))^{-1}. \quad (5.7)$$

Lemma 2.4 in [BGBHR10], which is an adaptation from [Via09] and [BMTY05], gives the derivative of the flow at a given time with respect to a variation  $(\varepsilon, t) \mapsto \xi(t) + \varepsilon \delta \xi(t) \in \mathfrak{g}$  of a smooth curve  $\xi = \xi_0$ . Namely,

$$\delta g_{t,s}^\xi := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} g_{t,s}^{\xi_\varepsilon} = g_{t,s}^\xi \int_s^t \left( \text{Ad}_{g_{s,r}^\xi} \delta \xi(r) \right) dr \in T_{g_{t,s}^\xi} G. \quad (5.8)$$

Importantly, formula (5.8) also holds for the diffeomorphisms case in a non-smooth setting, as shown in [TY05], where the assumption is  $\xi \in L^2([0, t_l], V)$ . Moreover, this proof can be adapted to the case of a finite dimensional Lie group. In particular, formula (5.8) can be used for  $\xi \in \mathcal{P}_{k-1}$ , whether one works with finite dimensional Lie groups or diffeomorphism groups.

Formula (5.8) and equation (5.6) are the key ingredients needed in order to take variations of the similarity measure in (5.2). With these preparations it is now straightforward to adapt the calculations done in the proof of Theorem 2.5 of [BGBHR10] to our case, in order to show that the following theorem holds.

**Theorem 5.1.** *A curve  $\xi \in \mathcal{P}_{k-1}$  is an extremal for the functional  $E$ , i.e.,  $\delta E = 0$  if and only if (I), (II), and (III) below hold:*

(I) *For  $t$  in any of the open intervals  $(0, t_1), \dots, (t_{l-1}, t_l)$ ,*

$$\sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}} = - \sum_{i=1}^l \chi_{[0, t_i]}(t) \left( g_{t,0}^\xi T_0 \diamond g_{t, t_i}^\xi \pi^i \right), \quad (5.9)$$

where  $\pi^i$  is defined by

$$\pi^i := \frac{1}{\sigma^2} \left( g_{t_i, 0}^\xi T_0 - I_{t_i} \right)^\flat \in V^*,$$

and  $\chi_{[0, t_i]}$  is the characteristic function of the interval  $[0, t_i]$ .

(II) For  $i = 1, \dots, l-1$  and  $r = 0, \dots, k-2$ ,

$$\lim_{t \rightarrow t_i^-} \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t) = \lim_{t \rightarrow t_i^+} \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t). \quad (5.10)$$

(III) For  $r = 0, \dots, k-2$ ,

$$\sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l) = 0. \quad (5.11)$$

Note that there is no condition at  $t_0 = 0$  analogous to (III) because of the fixed end point conditions  $\xi^{(j)}(0) = \xi_0^j$ , for  $j = 0, \dots, k-2$ .

*Proof.* Set  $t_0 = 0$  for convenience. A series of partial integrations taking into account the fixed end point conditions  $\xi^{(j)}(0) = \xi_0^j$ ,  $j = 0, \dots, k-2$ , leads to

$$\begin{aligned} \delta \int_0^{t_l} \ell dt &= \sum_{i=0}^{l-1} \int_{t_i}^{t_{i+1}} \sum_{j=0}^{k-1} \left\langle \frac{\delta \ell}{\delta \xi^{(j)}}, \delta \xi^{(j)} \right\rangle dt \\ &= \sum_{i=0}^{l-1} \int_{t_i}^{t_{i+1}} \left\langle \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}}(t), \delta \xi(t) \right\rangle dt \\ &\quad + \sum_{i=1}^{l-1} \sum_{r=0}^{k-2} \left( \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \left( \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^-) - \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^+) \right), \delta \xi^{(r)}(t_i) \right\rangle \right. \\ &\quad \left. + \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l), \delta \xi^{(r)}(t_l) \right\rangle \right). \end{aligned} \quad (5.12)$$

Note that the hypothesis  $\xi \in \mathcal{P}_{k-1}$  is sufficient to give meaning to the previous formula.

On the other hand, using formula (5.8) and mimicking the computations done in [BGBHR10], one finds for the variation of the similarity measure

$$\delta \left( \frac{1}{2\sigma^2} \sum_{i=1}^l \|g^\xi(t_i)T_0 - I_{t_i}\|_V^2 \right) = \int_0^{t_l} \left\langle \sum_{i=1}^l \chi_{[0,t_i]}(t) \left( g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right), \delta \xi(t) \right\rangle dt. \quad (5.13)$$

Assembling the two contributions to  $\delta E$ , we arrive at

$$\begin{aligned} \delta E &= \sum_{s=0}^{l-1} \int_{t_s}^{t_{s+1}} \left\langle \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}}(t) + \sum_{i=1}^l \chi_{[0,t_i]}(t) \left( g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right), \delta \xi(t) \right\rangle dt \\ &\quad + \sum_{i=1}^{l-1} \sum_{r=0}^{k-2} \left( \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \left( \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^-) - \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_i^+) \right), \delta \xi^{(r)}(t_i) \right\rangle \right. \\ &\quad \left. + \sum_{r=0}^{k-2} \left\langle \sum_{j \geq r+1}^{k-1} (-1)^{j-r-1} \frac{d^{j-r-1}}{dt^{j-r-1}} \frac{\delta \ell}{\delta \xi^{(j)}}(t_l), \delta \xi^{(r)}(t_l) \right\rangle \right). \end{aligned} \quad (5.14)$$

Stationarity  $\delta E = 0$  therefore leads to equations (5.9) - (5.11).  $\square$

**Remark 5.2.** The right-hand side of equation (5.9) follows coadjoint motion on every open interval  $(0, t_1), \dots, (t_l, t_{l+1})$ . Therefore,

$$\left( \frac{d}{dt} + \text{ad}_{\xi(t)}^* \right) \sum_{j=0}^{k-1} (-1)^j \frac{d^j}{dt^j} \frac{\delta \ell}{\delta \xi^{(j)}} = 0, \quad (5.15)$$

in which we once again recognize the higher-order Euler-Poincaré equation (3.4).

### 5.3 Two examples of interest for computational anatomy

Regarding potential applications in *CA*, an interesting property of higher-order models is the gain in smoothness of the optimal path  $T : t \in [0, t_l] \mapsto g^\xi(t)T_0 \in V$ , in comparison with first-order models. For instance, in the case of piecewise-geodesic (i.e., first-order) interpolation, where  $\ell(\xi) := \frac{1}{2}\|\xi\|_{\mathfrak{g}}^2$ , equation (5.9) reads

$$\xi(t) = - \sum_{i=1}^l \chi_{[0, t_i]}(t) \left( g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right). \quad (5.16)$$

In general therefore,  $\xi$  will be discontinuous at each time point  $t_i$  for  $i < l$ , which implies non-differentiability of  $T$  at these points. In contrast, for the Lagrangian  $\ell_1(\dot{\xi}) := \frac{1}{2}\|\dot{\xi}\|_{\mathfrak{g}}^2$ , equation (5.9) becomes

$$\ddot{\xi}(t) = \sum_{i=1}^l \chi_{[0, t_i]}(t) \left( g_{t,0}^\xi T_0 \diamond g_{t,t_i}^\xi \pi^i \right). \quad (5.17)$$

Now the curves  $\xi(t)$  and  $T(t)$  are  $C^1$  and  $C^2$  on  $[0, t_l]$ , respectively. Note that the inexact interpolation we consider here yields a  $C^2$  curve  $T$ , whereas the exact interpolation method presented in the example of Section 2.2 leads to  $C^1$  solutions. Note also that the minimization of the functional  $E$  for  $\ell_1$  when  $l = 1$  produces Lie-exponential solutions on  $G$ . More precisely, if the Lie-exponential map is surjective and the action of  $G$  on  $V$  is transitive, then there exists  $\xi_0 \in \mathfrak{g}$  such that  $\exp(t_1 \xi_0)T_0 = I_{t_1}$ . Hence, the constant curve  $\xi \equiv \xi_0$  is a minimizer of the functional  $E$ , with  $E[\xi] = 0$ . The Lie-exponential has been widely used in *CA*, for instance in [AFPA06, Ash07].

Another Lagrangian of interest for *CA* is  $\ell_2(\xi, \dot{\xi}) := \frac{1}{2}\|\dot{\xi} + \text{ad}_{\xi}^\dagger \xi\|_{\mathfrak{g}}^2$ , which measures the acceleration on the Lie group for the right-invariant metric induced by the norm  $\|\cdot\|_{\mathfrak{g}}$ . The Lagrangian  $\ell_2$  may therefore have more geometrical meaning than  $\ell_1$ . However,  $\ell_1$  is worth studying since it is simpler from both the computational and the analytical point of view: The existence of a minimizer for  $\ell_1$  can be obtained straightforwardly following the strategy of [TY05]. In contrast, a deeper analytical study is required for  $\ell_2$ , since analytical issues arise in infinite dimensions.

### 5.4 Template matching on the sphere

Consider as a finite-dimensional example  $G = SO(3)$  with norm  $\|\Omega\|_{\mathfrak{so}(3)} = \sqrt{\Omega \cdot \mathbf{I} \Omega}$  on the Lie algebra  $\mathfrak{so}(3)$ , where  $\mathbf{I}$  is a symmetric positive-definite matrix (the moment of inertia tensor). Let  $V = \mathbb{R}^3$  with  $\|\cdot\|_{\mathbb{R}^3}$  the Euclidean distance. We would like to interpolate a time sequence of points on the unit sphere  $S^2 \subset \mathbb{R}^3$  starting from the template  $T_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . Choose the times to be

$t_i = \frac{1}{5}i$  for  $i = 1, \dots, 5$ , and

$$I_{t_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad I_{t_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad I_{t_3} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad I_{t_4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad I_{t_5} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \quad (5.18)$$

The associated minimization problem for a given Lagrangian  $\ell(\mathbf{\Omega}, \dots, \mathbf{\Omega}^{(k-1)})$  is:

Minimize

$$E[\mathbf{\Omega}] := \int_0^1 \ell(\mathbf{\Omega}(t), \dots, \mathbf{\Omega}^{(k-1)}(t)) dt + \frac{1}{2\sigma^2} \sum_{i=1}^5 \left\| \Lambda^{\mathbf{\Omega}}(t_i) T_0 - I_{t_i} \right\|_{\mathbb{R}^3}^2, \quad (5.19)$$

subject to the conditions  $\mathbf{\Omega}^{(j)}(0) = \mathbf{\Omega}_0^j$ ,  $j = 0, \dots, k-2$ , where  $\Lambda^{\mathbf{\Omega}}(t)$  is a continuous curve defined by

$$\Lambda^{\mathbf{\Omega}}(0) = e, \quad \text{and} \quad \frac{d}{dt} \Lambda^{\mathbf{\Omega}}(t) = \Omega(t) \Lambda^{\mathbf{\Omega}}(t),$$

whenever  $t$  is in one of the open intervals  $(0, t_1), \dots, (t_4, t_5)$ . As we mentioned in Section 5.3, an important property of higher-order models is the increase in smoothness of the optimal path when compared with first-order models. We illustrate this behavior in Figures 5.1 and 5.2:

Figure 5.1 shows the interpolation between the given points  $I_{t_1}, \dots, I_{t_5}$  for the first order Lagrangian

$$\ell(\mathbf{\Omega}) = \frac{1}{2} \mathbf{\Omega} \cdot \mathbf{I} \mathbf{\Omega}. \quad (5.20)$$

We contrast this with the second order model

$$\ell(\mathbf{\Omega}, \dot{\mathbf{\Omega}}) = \frac{1}{2} \left( \dot{\mathbf{\Omega}} + \mathbf{I}^{-1}(\mathbf{\Omega} \times \mathbf{I} \mathbf{\Omega}) \right) \cdot \mathbf{I} \left( \dot{\mathbf{\Omega}} + \mathbf{I}^{-1}(\mathbf{\Omega} \times \mathbf{I} \mathbf{\Omega}) \right). \quad (5.21)$$

Note that this is the reduced Lagrangian for splines on  $SO(3)$ , as we discussed in Section 3, and for  $\mathbf{I} = e$  we recognize equation (5.15) to be the NHP equation (3.22).

Figure 5.2 visualizes the resulting interpolation for two different choices of the moment of inertia tensor  $\mathbf{I}$ , namely

$$\mathbf{I}_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{I}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.22)$$

In order to compare the two cases we have normalized  $\mathbf{I}_2$  in such a way that it has the same norm as  $\mathbf{I}_1$  with respect to the norm  $\|\mathbf{I}\|^2 = \text{tr}(\mathbf{I}^T \mathbf{I})$ . The figures were obtained by minimizing the functional  $E$  using the downhill simplex algorithm `fmin_tnc` that is included in the *optimize* package of SciPy, [JOP<sup>+</sup>].

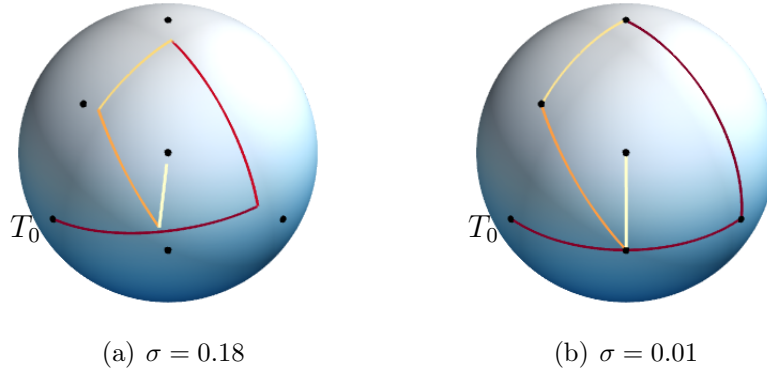


Fig. 5.1: First order template matching results are shown for the Lagrangian (5.20) with  $\mathbf{I} = e$ , for two different values of tolerance  $\sigma$ . These values have been chosen so that the sum of the mismatch penalties is similar in size to the one obtained in the second order template matching shown in Figure 5.2. As might be expected, when the tolerance is smaller, the first order curves pass nearer their intended target points. These first order curves possess jumps in tangent directions at the beginning of each new time interval.

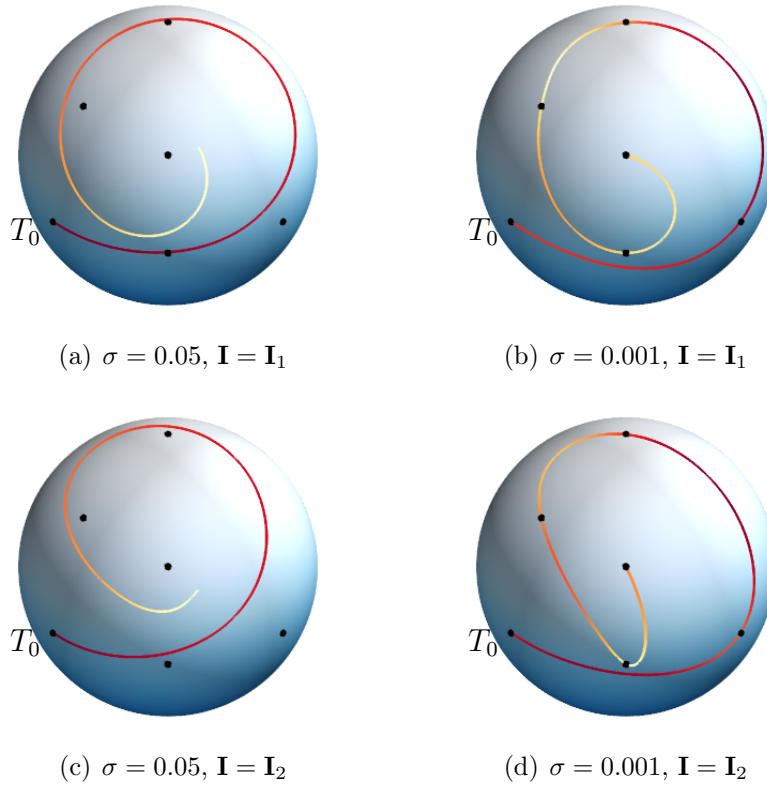


Fig. 5.2: The pictures in the top row show the template matching for the Lagrangian (5.20) with  $\mathbf{I}_1$  with two different values of tolerance,  $\sigma$ . The bottom row represents the corresponding matching results for  $\mathbf{I}_2$ . One observes that the quality of matching increases as the tolerance decreases. This is due to the increased weight on the penalty term in (5.2). The color of the curves represents the magnitude of the velocity vector of the curve on the sphere (red is large, white is small). We fixed the initial angular velocity  $\boldsymbol{\Omega}(0) = \frac{5\pi}{2} (0, 0, 1)$ . On comparing these figures with those in the first order case, one observes that the second-order method produces smoother curves.



**Remark 5.3.** Standard variational calculus arguments ensure the existence of a minimizer to the functional (5.2) with Lagrangian (5.21). In Theorem 5.1, we chose to fix  $\xi^{(j)}(0) = \xi_0^j$ ,  $j = 0, \dots, k-2$ , which reduces in this case to fixing  $\Omega(0)$ . We might, however, also want to optimize over this initial velocity. Unfortunately, examples can be exhibited where there does not exist any solution to the minimization of  $E$  if one also minimizes over  $\Omega(0)$ . One possibility to restore well-posedness while retaining the minimization over  $\Omega(0)$  is to modify  $E$  by adding a penalty on the norm  $\|\Omega(0)\|$ .

The situation in infinite dimensions is similar, however proving existence results would require much deeper analytical study than in the finite dimensional case.

In this section we presented higher-order methods that increase the smoothness in interpolating through a sequence of data points. In future work these methods will be compared to the shape spline model introduced in [TV10]. Also of interest for  $CA$  is the metamorphosis approach that is discussed briefly in Section 6.

## 6 Optimization with penalty

This section adapts the optimization approach of [GBHR10] to higher-order Lagrangians. As in the case of the Clebsch-Pontryagin approach, one considers the (right or left) action  $\Phi : G \times Q \rightarrow Q$  of a Lie group  $G$  on a manifold  $Q$ . The basic idea is to replace the constraints in the Clebsch optimal control problem with a penalty function added to the cost function and to obtain in this way a classical (unconstrained) optimization problem. The penalty term is expressed with the help of a Riemannian metric  $\gamma$  on the manifold  $Q$ . Given a cost function  $\ell : k\mathfrak{g} \times Q \rightarrow \mathbb{R}$ ,  $\sigma > 0$  and the elements  $n_0, n_T \in Q$ ,  $\xi_0^j, \xi_T^j \in \mathfrak{g}$ ,  $j = 0, \dots, k-2$ , one minimizes

$$\int_0^T \left( \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}, n) + \frac{1}{2\sigma^2} \|\dot{n} - \xi_Q(n)\|^2 \right) dt, \quad (6.1)$$

over curves  $t \mapsto n(t) \in Q$  and  $t \mapsto \xi(t) \in \mathfrak{g}$  such that

$$n(0) = n_0, \quad n(T) = n_T, \quad \xi^{(j)}(0) = \xi_0^j, \quad \xi^{(j)}(T) = \xi_T^j, \quad j = 0, \dots, k-2,$$

where  $\|\cdot\|$  is the norm on  $TQ$  induced by the metric  $\gamma$  and, as in the Clebsch-Pontryagin case in Section 4,  $\xi_Q(n)$  denotes the infinitesimal generator of the  $G$ -action associated to  $\xi \in \mathfrak{g}$ , evaluated at  $n \in Q$ . The corresponding stationarity conditions are found to be:

$$\sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = \mathbf{J}(\pi), \quad \dot{n} = \xi_Q(n) + \sigma^2 \pi^\sharp, \quad \frac{D}{Dt} \pi = -\langle \pi, \nabla \xi_Q \rangle + \frac{\partial \ell}{\partial n}, \quad (6.2)$$

where the notation

$$\pi := \frac{1}{\sigma^2} \nu_n^\flat = \frac{1}{\sigma^2} (\dot{n} - \xi_Q(n))^\flat \in T^*Q$$

has been used and the covariant derivatives  $D/Dt$  and  $\nabla$  are associated to the Riemannian metric  $\gamma$  on  $Q$ .

These equations should be compared with the stationarity conditions (4.2) associated to the Clebsch approach, which can be rewritten, with the help of a Riemannian metric, as

$$\sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = \mathbf{J}(\alpha), \quad \dot{q} = \xi_Q(q), \quad \frac{D}{Dt} \alpha = -\langle \alpha, \nabla \xi_Q \rangle + \frac{\partial \ell}{\partial q}. \quad (6.3)$$

Before proceeding further, we will pause to define some additional notation that will be convenient later.

**Definition 6.1.** Consider a Lie group  $G$  acting on a Riemannian manifold  $(Q, \gamma)$ . We define the  $\mathfrak{g}^*$ -valued  $(1, 1)$  tensor field  $\mathcal{F}^\nabla : T^*Q \times TQ \rightarrow \mathfrak{g}^*$  associated to the Levi-Civita connection  $\nabla$  by

$$\langle \mathcal{F}^\nabla(\alpha_q, u_q), \eta \rangle := \langle \alpha_q, \nabla_{u_q} \eta(q) \rangle, \quad (6.4)$$

for all  $u_q \in T_q Q$ ,  $\alpha_q \in T_q^* Q$ , and  $\eta \in \mathfrak{g}$ .

The main properties of the tensor field  $\mathcal{F}^\nabla$  are discussed in [GBHR10], where one also finds the proofs of the following two lemmas about the properties of  $\mathcal{F}^\nabla$ . The first lemma below relates  $\mathcal{F}^\nabla$  to the connectors of the covariant derivatives on  $TQ$  and  $T^*Q$ . The second lemma explains that  $\mathcal{F}^\nabla$  is antisymmetric under transposition in the inner product defined by the Riemannian metric  $\gamma$  when  $G$  acts by isometries.

**Lemma 6.2.** For all  $\alpha_q \in T_q^* Q$ ,  $u_q \in T_q Q$ , and  $\xi \in \mathfrak{g}$ ,

$$\langle \mathcal{F}^\nabla(\alpha_q, u_q), \xi \rangle = \langle \alpha_q, K(\xi_{TQ}(u_q)) \rangle = -\langle K(\xi_{T^*Q}(\alpha_q)), u_q \rangle,$$

where  $K$  denotes the connectors of the covariant derivatives on  $TQ$  and  $T^*Q$ , respectively.

*Proof.* See the proof of Lemma 3.5 in [GBHR10, §3].  $\square$

**Remark 6.3.** A detailed treatment of connectors and their associated linear connections for covariant derivatives can be found in [Mic08, §13.8]. We also refer to [GBHR10] for useful properties of the connector  $K$  of relevance to the present paper. In infinite dimensions one needs to assume that the given weak Riemannian metric has a smooth geodesic spray  $S \in \mathfrak{X}(TG)$ , but such analytical issues will not be of concern to us here.

**Lemma 6.4.** If  $G$  acts by isometries, then  $\mathcal{F}^\nabla$  is antisymmetric, that is

$$\mathcal{F}^\nabla(\alpha_q, u_q) = -\mathcal{F}^\nabla(u_q^\flat, \alpha_q^\sharp),$$

for all  $u_q \in T_q Q$ ,  $\alpha_q \in T_q^* Q$ .

*Proof.* Since  $G$  acts by isometries,  $\mathcal{L}_{\xi_Q} g = 0$ ; which implies  $(\nabla \xi_Q)^T = -\nabla \xi_Q$ .  $\square$

The tensor field  $\mathcal{F}^\nabla$  arises naturally in computing the equations of motion associated to the stationarity conditions (6.2) for optimization with penalty. A computation, similar to the one given in [GBHR10, §3] in the first order case, yields

$$\begin{cases} (\partial_t \mp \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell}{\delta \xi^{(j)}} = \mathbf{J} \left( \frac{\partial \ell}{\partial n} \right) + \frac{1}{\sigma^2} \mathcal{F}^\nabla(\nu_n^\flat, \nu_n), \\ \frac{D}{Dt} \nu_n^\flat = -\langle \nu_n^\flat, \nabla \xi_Q \rangle + \sigma^2 \frac{\partial \ell}{\partial n}, \quad \nu_n := \dot{n} - \xi_Q(n), \end{cases} \quad (6.5)$$

where in  $(\mp)$  one chooses  $-$  (resp.  $+$ ) when  $G$  acts on  $Q$  by a right (resp. left) action, consistently with (4.5).

As a consequence of Lemma 6.4, if  $G$  acts on  $(Q, \gamma)$  by isometries, then the term  $\frac{1}{\sigma^2} \mathcal{F}^\nabla(\nu_n^\flat, \nu_n)$  vanishes so that the optimization problem (6.1) produces the  $k^{\text{th}}$  order Euler-Poincaré equations.

**Example: The NHP equation via optimization.** In this case, since  $SO(3)$  acts by isometries on  $\mathbb{R}^3$ , the minimization problem

$$\min \int_{t_0}^{t_1} \left( \frac{1}{2} \|\dot{\mathbf{q}}\|^2 + \frac{1}{2\sigma^2} \|\dot{\mathbf{q}} - \mathbf{\Omega} \times \mathbf{q}\|^2 \right) dt$$

produces the NHP equations (1.1).

**Metamorphosis and Lagrange-Poincaré reductions.** Equations (6.5) may also be obtained by a generalization of the metamorphosis reduction developed in [GBHR10], as follows. For simplicity, we only treat the case of a right action of  $G$  on  $Q$ .

Consider a  $G$ -invariant Lagrangian  $L = L(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}) : T^{(k)}G \times TQ \rightarrow \mathbb{R}$  relative to the action of  $h \in G$  given by

$$(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}) \mapsto (hg, h\dot{g}, \dots, hg^{(k)}, qh^{-1}, \dot{q}h^{-1})$$

and consider the quotient map

$$(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}) \mapsto (\xi, \dot{\xi}, \dots, \xi^{(k-1)}, n, \nu) \in k\mathfrak{g} \times TQ, \quad \xi = g^{-1}\dot{g}, \quad n = qg, \quad \nu = \dot{q}g. \quad (6.6)$$

The equations of motion for the reduced Lagrangian  $\ell_M$  induced by  $L$  on  $k\mathfrak{g} \times TQ$  can be obtained by a direct generalization of the method used in [GBHR10] for  $k = 1$ . If  $L$  has the particular form

$$L(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}) = \mathcal{L}(g, \dot{g}, \dots, g^{(k)}, q) + \frac{1}{2\sigma^2} \|\dot{q}g\|^2,$$

where  $\mathcal{L}$  is the  $G$ -invariant Lagrangian associated to the function  $\ell$  in (6.1), then we recover equations (6.5) (with the upper sign chosen).

Instead of the so-called metamorphosis quotient map (6.6) one may also use Lagrange-Poincaré reduction with the quotient map

$$(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}) \mapsto (\xi, \dot{\xi}, \dots, \xi^{(k-1)}, n, \dot{n}) \in k\mathfrak{g} \times TQ, \quad \xi = g^{-1}\dot{g}, \quad n = qg. \quad (6.7)$$

The reduced equations of motion for metamorphosis with geometric splines that arise in the Lagrange-Poincaré approach are

$$\begin{cases} \frac{\delta \ell_{LP}}{\delta n} - \frac{d}{dt} \frac{\delta \ell_{LP}}{\delta \dot{n}} = 0, \\ (\partial_t \mp \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell_{LP}}{\delta \xi^{(j)}} = 0, \end{cases} \quad (6.8)$$

(with the upper sign chosen) where  $\ell_{LP}$  is the reduced Lagrangian associated to the same unreduced Lagrangian  $L$  as before, but using the quotient map (6.7) instead of (6.6).

Note that the Lagrange-Poincaré approach generalizes easily to higher-order Lagrangians in  $q$  such as  $L := L(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}, \dots, q^{(k)}) : T^{(k)}(G \times Q) \rightarrow \mathbb{R}$ . The equations of motions are then simply

$$\begin{cases} \sum_{j=0}^{k-1} (-1)^j \frac{\delta \ell_{LP}}{\delta n^{(j)}} = 0, \\ (\partial_t \pm \text{ad}_\xi^*) \sum_{j=0}^{k-1} (-1)^j \partial_t^j \frac{\delta \ell_{LP}}{\delta \xi^{(j)}} = 0. \end{cases} \quad (6.9)$$

The metamorphosis reduction approach also generalizes to higher higher-order Lagrangians in  $q$ . In this case, one uses the quotient map

$$(g, \dot{g}, \dots, g^{(k)}, q, \dot{q}, \dots, q^{(k)}) \mapsto (\xi, \dot{\xi}, \dots, \xi^{(k-1)}, n, \nu_1, \dots, \nu_k) \in k\mathfrak{g} \times T^{(k)}Q, \quad (6.10)$$

where  $\xi = g^{-1}\dot{g}$  and  $(n, \nu_1, \dots, \nu_k) = \Phi_g^{(k)}(q, \dot{q}, \dots, q^{(k)})$ ,  $\Phi^{(k)}$  being the natural induced action of  $G$  on  $T^{(k)}Q$ . However for  $k \geq 2$ , the associated reduced equations are quite complex on general Riemannian manifolds so one may prefer to use the equivalent Lagrange-Poincaré formulation (6.9).

**Remark 6.5.** The idea of metamorphosis with splines may apply in imaging as in [HTY09] by using, e.g.,  $L(g_t, \dot{g}_t, \ddot{g}_t, \eta_t, \dot{\eta}_t)$ ,  $L(g_t, \dot{g}_t, \eta_t, \dot{\eta}_t, \ddot{\eta}_t)$ , or  $L(g_t, \dot{g}_t, \ddot{g}_t, \eta_t, \dot{\eta}_t, \ddot{\eta}_t)$ , instead of  $L(g_t, \dot{g}_t, \eta_t, \dot{\eta}_t)$ .

## 7 Clebsch and Lie-Poisson-Ostrogradsky formulations

In this Section we present two Hamiltonian formulations associated to the higher order Euler-Poincaré equations (3.29) with  $q$ -dependence. (The case of equations (3.30) may be obtained by making obvious modifications.) The first is a canonical Hamiltonian formulation that generalizes to higher order the canonical Clebsch formulation of Euler-Poincaré dynamics. The second is a generalization of the Lie-Poisson formulation (with  $q$ -dependence) to higher order, that uses Ostrogradsky momenta. We now recall these formulations in the first order case.

**Clebsch canonical formulation.** This is associated to the optimal control formulation described in §4. In the case  $k = 1$  the canonical Hamiltonian formulation is already given by the Pontryagin approach. Indeed, if  $\xi \mapsto \frac{\delta \ell}{\delta \xi}$  is a diffeomorphism we consider the function  $h : \mathfrak{g} \times Q \rightarrow \mathbb{R}$  defined by

$$h(\mu, q) := \langle \mu, \xi \rangle - \ell(\xi, q), \quad \frac{\delta \ell}{\delta \xi} = \mu$$

and the collective Hamiltonian  $H : T^*Q \rightarrow \mathbb{R}$  given by  $H(\alpha_q) := h(\mathbf{J}(\alpha_q), q)$ . If  $\alpha_q(t)$  is a solution of Hamilton's canonical equations for  $H$  on  $T^*Q$ , then  $(\mu(t), q(t))$ , where  $\mu(t) := \mathbf{J}(\alpha_q(t))$ , is a solution of the Euler-Poincaré equations

$$(\partial_t \pm \text{ad}_\xi^*) \frac{\delta \ell}{\delta \xi} = \mathbf{J} \left( \frac{\delta \ell}{\delta q} \right). \quad (7.1)$$

This canonical formulation of the Euler-Poincaré equations recovers some important examples such as the Clebsch variables for the ideal fluid [MW83], singular solutions of the Camassa-Holm equations [HM04], and double bracket equations, as explained in [GBR09].

**Lie-Poisson formulation.** This is obtained by reduction of the Hamiltonian  $H_{q_0} : T^*G \rightarrow \mathbb{R}$  associated to  $L_{q_0}$  by Legendre transformation. If  $L$  is  $G$ -invariant as a function defined on  $TG \times Q$ , then  $H : T^*G \times Q \rightarrow \mathbb{R}$  is  $G$ -invariant and therefore induces the Hamiltonian  $h$  given above. By Poisson reduction of the manifold  $T^*G \times Q$ , where  $Q$  is endowed with the trivial Poisson structure, one obtains the Lie-Poisson equations

$$\left( \partial_t \pm \text{ad}_{\frac{\delta h}{\delta \mu}}^* \right) \mu = -\mathbf{J} \left( \frac{\delta h}{\delta q} \right), \quad \partial_t q - \left( \frac{\delta h}{\delta \mu} \right)_Q(q) = 0,$$

together with the associated Poisson structure

$$\{f, g\}(\mu, q) = \pm \left\langle \mu, \left[ \frac{\delta f}{\delta \mu}, \frac{\delta g}{\delta \mu} \right] \right\rangle + \left\langle \mathbf{J} \left( \frac{\delta f}{\delta q} \right), \frac{\delta g}{\delta \mu} \right\rangle - \left\langle \mathbf{J} \left( \frac{\delta g}{\delta q} \right), \frac{\delta f}{\delta \mu} \right\rangle \quad (7.2)$$

on  $\mathfrak{g}^* \times Q$ ; see [GBT10]. These equations are equivalent to their Lagrangian counterpart (7.1).

## 7.1 Higher order Clebsch formulations

### Second order case

Recall from Definition 4.1 that the Clebsch-Pontryagin variational formulation of the second order Euler-Poincaré equations reads

$$\delta \int_{t_0}^{t_1} \left( \ell(\xi, \dot{\xi}, q) + \langle \alpha, \dot{q} - \xi_Q(q) \rangle \right) dt = 0,$$

over curves  $\xi(t) \in \mathfrak{g}$  and  $\alpha(t) \in T_{q(t)}^*Q$  and under conditions (A), (B), (C). If  $\dot{\xi} \mapsto \pi := \delta \ell / \delta \dot{\xi}$  is a diffeomorphism, we define  $h(\xi, \pi, q) := \langle \pi, \dot{\xi} \rangle - \ell(\xi, \dot{\xi}, q)$  and the Pontryagin variational principle may be written equivalently as

$$\delta \int_{t_0}^{t_1} \left( \langle \pi, \dot{\xi} \rangle - h(\xi, \pi, q) + \langle \alpha, \dot{q} - \xi_Q(q) \rangle \right) dt = 0, \quad \text{where } \pi := \frac{\delta \ell}{\delta \dot{\xi}}$$

over curves  $\xi(t) \in \mathfrak{g}$  and  $\alpha(t) \in T_{q(t)}^*Q$ . Equivalently, this can be reformulated as

$$\delta \int_{t_0}^{t_1} \left( \langle \pi, \dot{\xi} \rangle - h(\xi, \pi, q) + \langle \alpha, \dot{q} - \xi_Q(q) \rangle \right) dt = 0,$$

over curves  $\xi(t) \in \mathfrak{g}$ ,  $\pi(t) \in \mathfrak{g}^*$ , and  $\alpha(t) \in T_{q(t)}^*Q$ , where  $\pi(t)$  is now an independent curve. The relation  $\pi = \delta \ell / \delta \dot{\xi}$  is recovered by variations of  $\mu(t)$ . One observes that this is simply the usual Hamilton Phase Space Variational Principle (i.e., not a Pontryagin Maximum Principle) on the phase space  $T^*(Q \times \mathfrak{g})$

$$\delta \int_0^1 H(\alpha, \xi, \pi) - \langle (\alpha, \pi), (\dot{q}, \dot{\xi}) \rangle dt = 0,$$

for the Hamiltonian

$$H : T^*(Q \times \mathfrak{g}) \rightarrow \mathbb{R}, \quad H(\alpha_q, \xi, \pi) := h(\xi, \pi, q) + \langle \mathbf{J}(\alpha_q), \xi \rangle.$$

We thus have proved the following result.

**Theorem 7.1.** *Let  $\ell : 2\mathfrak{g} \times Q \rightarrow \mathbb{R}$ ,  $\ell = \ell(\xi, \dot{\xi}, q)$  be a cost function such that  $\dot{\xi} \mapsto \pi := \delta \ell / \delta \dot{\xi}$  is a diffeomorphism and define the function*

$$h(\xi, \pi, q) := \langle \pi, \dot{\xi} \rangle - \ell(\xi, \dot{\xi}, q). \quad (7.3)$$

*Then the stationarity conditions (4.4) for the 2<sup>nd</sup> order Clebsch-Pontryagin optimal control problem (4.1) with cost function  $\ell$  are given by the canonical Hamilton equations on  $T^*(Q \times \mathfrak{g})$  relative to the Hamiltonian  $H(\xi, \mu, \alpha) = h(\xi, \pi, q) + \langle \mathbf{J}(\alpha), \xi \rangle$ .*

One can alternatively prove this result by computing explicitly the canonical Hamilton equations for  $H$  on  $T^*(Q \times \mathfrak{g})$ . We obtain

$$\dot{\alpha} = X_H(\alpha) = \xi_{T^*Q}(\alpha) - \text{Ver}_\alpha \frac{\delta h}{\delta q}, \quad \dot{\xi} = \frac{\delta H}{\delta \pi} = \frac{\delta h}{\delta \pi}, \quad \dot{\pi} = -\frac{\delta H}{\delta \xi} = \frac{\delta \ell}{\delta \xi} - \mathbf{J}(\alpha). \quad (7.4)$$

Clearly, these equations coincide with the stationarity conditions (4.2). In particular, the last equation reads

$$\mathbf{J}(\alpha) = \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}}.$$

### Example 1: Geodesic 2-spline equation on Lie groups

Recall from §3.2 that the reduced Lagrangian for 2-splines on a Lie group  $G$  with right  $G$ -invariant Riemannian metric reads

$$\ell(\xi, \dot{\xi}) = \frac{1}{2} \|\eta\|^2 = \frac{1}{2} \left\| \dot{\xi} + \text{ad}_\xi^\dagger \xi \right\|^2. \quad (7.5)$$

Here we denote

$$\eta = \dot{\xi} + \text{ad}_\xi^\dagger \xi \quad \text{with} \quad \text{ad}_\xi^\dagger \nu = (\text{ad}_\xi^*(\nu^\flat))^\sharp \quad \text{for} \quad \xi, \nu \in \mathfrak{g}. \quad (7.6)$$

Then the quantity computed in equation (3.20)

$$\mu := \frac{\delta \ell}{\delta \dot{\xi}} - \partial_t \frac{\delta \ell}{\delta \xi} = (\partial_t \eta + \text{ad}_\eta \xi + \text{ad}_\eta^\dagger \xi)^\flat, \quad \text{with} \quad \eta = \dot{\xi} + \text{ad}_\xi^\dagger \xi,$$

satisfies the  $2^{\text{nd}}$ -order Euler-Poincaré equation,  $(\partial_t + \text{ad}_\xi^*)\mu = 0$ , which is also the geometric 2-spline equation of [CS95]. We now consider the canonical formulation of 2-splines.

**Hamiltonian formulation of the geodesic 2-spline equation on  $T^*(Q \times \mathfrak{g})$ .** As we have seen, the Clebsch-Pontryagin approach of Section 4 allows the geodesic 2-spline equation to be recast as a set of canonical Hamilton equations for a Hamiltonian  $H : T^*(Q \times \mathfrak{g}) \rightarrow \mathbb{R}$ . Note that in the case of 2-splines, the variable  $\pi$  is

$$\pi = \frac{\delta \ell}{\delta \dot{\xi}} = \dot{\xi}^\flat + \text{ad}_\xi^* \xi^\flat = \eta^\flat$$

which proves that  $\dot{\xi} \mapsto \frac{\delta \ell}{\delta \dot{\xi}}$  is a diffeomorphism. One thus obtains the Hamiltonian

$$\begin{aligned} H(\alpha, \xi, \pi) &= \langle \pi, \dot{\xi} \rangle - \ell(\xi, \dot{\xi}) + \langle \mathbf{J}(\alpha), \xi \rangle, & \pi &= \frac{\delta \ell}{\delta \dot{\xi}} \\ &= \frac{1}{2} \|\pi\|^2 - \left\langle \pi, \text{ad}_\xi^\dagger \xi \right\rangle + \langle \mathbf{J}(\alpha), \xi \rangle, \end{aligned} \quad (7.7)$$

where  $\|\cdot\|$  denotes the norm induced by  $\gamma$  on  $\mathfrak{g}^*$ . The canonical Hamiltonian formulation (7.4) now yields the dynamical system

$$\dot{\alpha} = \xi_{T^*Q}(\alpha), \quad \dot{\xi} = \pi^\sharp - \text{ad}_\xi^\dagger \xi, \quad \dot{\pi} = -\text{ad}_{\pi^\sharp}^* \xi^\flat - (\text{ad}_{\pi^\sharp} \xi)^\flat - \mathbf{J}(\alpha) \quad (7.8)$$

As we have proved above, the Euler-Poincaré equation  $(\partial_t + \text{ad}_\xi^*)\mu = 0$  is then established by noticing that  $\mu = \mathbf{J}(\alpha)$  is the cotangent-lift momentum map for the action of the Lie group  $G$  on

the manifold  $Q$  and that the  $\dot{\xi}$ -equation implies  $\pi^\sharp = \dot{\xi} + \text{ad}_\xi^\dagger \xi = \eta$ . Note that the solution for the momentum map  $\mu = \mathbf{J}(\alpha)$  may be obtained entirely from the canonical Hamilton equations, without explicitly solving the Euler-Poincaré equation. For a bi-invariant metric, one has  $\text{ad}_\xi^\dagger \xi = 0$  in the  $\dot{\xi}$ -equation and the last two terms cancel each other in the  $\dot{\pi}$ -equation. Consequently, these two canonical equations simplify to  $\dot{\xi} = \pi^\sharp$  and  $\dot{\pi} = -\mathbf{J}(\alpha)$ . From them, we find

$$\ddot{\xi} = -\mathbf{J}(\alpha)^\sharp \quad \text{and} \quad \ddot{\xi} = -\text{ad}_\xi^\dagger \ddot{\xi}, \quad (7.9)$$

in agreement with equation (3.21) and reference [CS95].

### Example 2: Geodesic 2-spline equations on $SO(3)$

We consider the particular case of the Lie group  $G = SO(3)$  endowed with the bi-invariant metric induced by the standard Ad-invariant inner product

$$\gamma(\Omega, \Gamma) = -\frac{1}{2} \text{Tr}(\Omega \Gamma).$$

We identify the dual  $\mathfrak{so}(3)^*$  with  $\mathfrak{so}(3)$  using  $\gamma$  so that  $\Omega^\flat = \Omega$ . Using the hat map  $\hat{\cdot}: \mathfrak{so}(3) \rightarrow \mathbb{R}^3$  (see (3.23)), the Euler-Poincaré equation in (7.9) reads

$$\ddot{\Omega} - \Omega \times \ddot{\Omega} = 0, \quad (7.10)$$

which was first found in [NHP89]. The difference in sign from that paper arises here from the choice of reduction by right-invariance instead of left-invariance.

**Canonical Hamilton equations on  $T^*\mathbb{R}^3 \times T^*\mathbb{R}^3$  for the NHP equation.** The Hamiltonian formulation of the NHP equation (7.10) for geometric splines on  $SO(3)$  with a bi-invariant metric may be obtained in canonical variables  $(\Omega, \pi, \mathbf{q}, \mathbf{p}) \in T^*\mathbb{R}^3 \times T^*\mathbb{R}^3$  from the Hamiltonian (see (7.7)),

$$H(\Omega, \pi, \mathbf{q}, \mathbf{p}) = \frac{1}{2} \|\pi\|^2 + \Omega \cdot \mathbf{q} \times \mathbf{p}. \quad (7.11)$$

This corresponds to the choice  $Q = \mathbb{R}^3$  on which  $SO(3)$  acts by matrix multiplication.

This Hamiltonian produces canonical equations of the form,

$$\begin{aligned} \dot{\mathbf{q}} &= \frac{\delta H}{\delta \mathbf{p}} = \Omega \times \mathbf{q}, & \dot{\mathbf{p}} &= -\frac{\delta H}{\delta \mathbf{q}} = \Omega \times \mathbf{p}, \\ \dot{\Omega} &= \frac{\delta H}{\delta \pi} = \pi, & \dot{\pi} &= -\frac{\delta H}{\delta \Omega} = -\mathbf{q} \times \mathbf{p}. \end{aligned}$$

The  $(\mathbf{q}, \mathbf{p})$ -equations here imply that  $\mu = \mathbf{J}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p}$  obeys the Euler-Poincaré equation for right invariance,

$$(\partial_t + \text{ad}_\Omega^*) \mu = 0 = \dot{\mu} - \Omega \times \mu,$$

which results in the NHP equation (7.10) when we substitute  $\mu = -\ddot{\Omega}$ .

The canonical Hamiltonian formulation of the NHP equation provides some insight into the interpretation of its constants of motion. For example, the Hamiltonian (7.11) Poisson commutes with  $|\mathbf{q}|^2$ ,  $|\mathbf{p}|^2$ ,  $(\mathbf{q} \cdot \mathbf{p})$ , and  $|\mathbf{q} \times \mathbf{p}|^2$ , although only the last of these Poisson commutes with all the others. The Hamiltonian (7.11) also Poisson commutes with the vector  $\mathbf{K} = \Omega \times \pi - \mathbf{q} \times \mathbf{p} \in$



$\mathbb{R}^3$ . Although the components of  $\mathbf{K}$  satisfy Poisson bracket relations  $\{K_1, K_2\} = K_3$  and cyclic permutations with each other, their sum of squares  $K^2 = K_1^2 + K_2^2 + K_3^2$  again Poisson commutes with all the others. The presence of the two constants of motion  $|\mathbf{q} \times \mathbf{p}|^2$  and  $K^2$  in Poisson involution allows symplectic reduction from six degrees of freedom to four, but the reduced system is still far from being integrable. The Hamiltonian conservation laws may be expressed in terms of  $(\boldsymbol{\Omega}, \dot{\boldsymbol{\Omega}}, \ddot{\boldsymbol{\Omega}}) \in T^{(2)}\mathbb{R}^3$  as

$$|\mathbf{q} \times \mathbf{p}|^2 = |\ddot{\boldsymbol{\Omega}}|^2, \quad \mathbf{K} = \boldsymbol{\Omega} \times \dot{\boldsymbol{\Omega}} + \ddot{\boldsymbol{\Omega}}, \quad K^2 = |\boldsymbol{\Omega} \times \dot{\boldsymbol{\Omega}}|^2 + 2(\boldsymbol{\Omega} \times \dot{\boldsymbol{\Omega}}) \cdot \ddot{\boldsymbol{\Omega}} + |\ddot{\boldsymbol{\Omega}}|^2.$$

All of these conservation laws were known in the literature, but had previously not been given a Hamiltonian interpretation. The Hamiltonian interpretation of the NHP equation (7.10) in this setting is that the rotations act on the cross product  $\mathbf{m} = \mathbf{q} \times \mathbf{p}$  diagonally in  $\mathbf{q}$  and  $\mathbf{p}$ , so that  $\dot{\mathbf{m}} = \boldsymbol{\Omega} \times \mathbf{m}$  for  $\dot{\mathbf{q}} = \boldsymbol{\Omega} \times \mathbf{q}$  and  $\dot{\mathbf{p}} = \boldsymbol{\Omega} \times \mathbf{p}$ . This is also the essence of the symmetric representation of rigid body motion discussed, e.g., in [BC96b].

### Higher order case

The canonical Clebsch formulation presented above can be adapted to higher order cost functions  $\ell = \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}, q)$  as follows. If  $\xi^{(k-1)} \mapsto \frac{\delta \ell}{\delta \xi^{(k-1)}}$  is a diffeomorphism, we define the function

$$\begin{aligned} h(\xi, \dot{\xi}, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k, q) \\ := \langle \pi_2, \dot{\xi} \rangle + \langle \pi_3, \ddot{\xi} \rangle + \dots + \langle \pi_k, \xi^{(k-1)} \rangle - \ell(\xi, \dot{\xi}, \dots, \xi^{(k-1)}, q), \quad \pi_k = \frac{\delta \ell}{\delta \xi^{(k-1)}} \end{aligned} \quad (7.12)$$

and we consider the Hamiltonian  $H : T^*(Q \times (k-1)\mathfrak{g}) \rightarrow \mathbb{R}$  given by

$$H(\alpha_q, \xi, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k) := h(\xi, \dot{\xi}, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k, q) + \langle \mathbf{J}(\alpha_q), \xi \rangle.$$

A straightforward computation shows that the canonical Hamilton equations on  $T^*(Q \times (k-1)\mathfrak{g})$  for  $H$  produce the stationarity condition (7.1) of the  $k^{\text{th}}$ -order Clebsch-Pontryagin optimal control with cost function  $\ell$  and therefore imply the  $k^{\text{th}}$ -order Euler-Poincaré equations (3.29). We thus obtain the generalization of Theorem 7.1 for  $k^{\text{th}}$ -order cost functions.

## 7.2 Ostrogradsky-Lie-Poisson reduction

The procedure of Lie-Poisson reduction of the Hamilton-Ostrogradsky theory parallels that for higher-order Euler-Poincaré reduction and produces a different Hamiltonian formulation of the higher-order dynamics that applies to  $k \geq 2$ . At first, we will discuss the Hamilton-Ostrogradsky approach for the higher-order Hamiltonian formulation based purely on Lie group reduction, i.e., without introducing the action of the Lie group  $G$  on the manifold  $Q$ . Then we will remark on how  $q$ -dependence may be easily incorporated.

**Second order.** Consider a  $G$ -invariant second order Lagrangian  $L : T^{(2)}G \rightarrow \mathbb{R}$ ,  $L = L(g, \dot{g}, \ddot{g})$ . The Ostrogradsky momenta are defined by the fiber derivatives,

$$p_1 = \frac{\partial L}{\partial \dot{g}} - \partial_t \frac{\partial L}{\partial \ddot{g}}, \quad p_2 = \frac{\partial L}{\partial \ddot{g}}$$

and produce the Legendre transform  $(g, \dot{g}, \ddot{g}, \ddot{\ddot{g}}) \in T^{(3)}G \mapsto (g, \dot{g}, p_1, p_2) \in T^*(TG)$ . We refer to [dLR85] for the intrinsic definition of the Legendre transform for higher order Lagrangians. See also [BC96a] for an application on  $SO(3)$ .

When the Legendre transform is a diffeomorphism, the corresponding Hamiltonian  $H : T^*(TG) \rightarrow \mathbb{R}$  is defined by

$$H(g, \dot{g}, p_1, p_2) := \langle p_1, \dot{g} \rangle + \langle p_2, \ddot{g} \rangle - L(g, \dot{g}, \ddot{g})$$

and the canonical Hamilton equations for  $H$  are equivalent to the  $2^{nd}$ -order Euler-Lagrange equations for  $L$ .

Applying reduction by symmetry to  $H$  induces a Hamiltonian  $h(\pi_1, \xi, \pi_2)$  on  $T^*(TG)/G \simeq \mathfrak{g}^* \times T^*\mathfrak{g}$ , which is related to the symmetry-reduced Lagrangian  $\ell(\xi, \dot{\xi})$  by the corresponding Legendre transformation,

$$h(\pi_1, \xi, \pi_2) = \left\langle \pi_1, \xi \right\rangle + \left\langle \pi_2, \dot{\xi} \right\rangle - \ell(\xi, \dot{\xi}), \quad \frac{\delta \ell}{\delta \dot{\xi}} = \pi_2. \quad (7.13)$$

By Reduction of the Hamilton-Ostrogradsky equations for  $H$  on  $T^*(TG)$  we obtain the *Ostrogradsky-Lie-Poisson* equations for  $h$

$$\begin{cases} \partial_t \pi_1 \pm \text{ad}_{\frac{\delta h}{\delta \pi_1}}^* \pi_1 = 0, \\ \partial_t \xi = \frac{\delta h}{\delta \pi_2}, \quad \partial_t \pi_2 = -\frac{\delta h}{\delta \xi}, \end{cases} \quad (7.14)$$

together with the non-canonical Poisson bracket given by

$$\begin{aligned} \{f, g\}(\pi_1, \xi, \pi_2) &= \pm \left\langle \pi_1, \left[ \frac{\delta f}{\delta \pi_1}, \frac{\delta g}{\delta \pi_1} \right] \right\rangle + \left\langle \frac{\delta f}{\delta \xi}, \frac{\delta g}{\delta \pi_2} \right\rangle - \left\langle \frac{\delta g}{\delta \xi}, \frac{\delta f}{\delta \pi_2} \right\rangle \\ &= \{f, g\}_{\pm}(\pi_1) + \{f, g\}_{can}(\xi, \pi_2) \end{aligned} \quad (7.15)$$

for functions  $f, g$  depending on the variables  $(\pi_1, \xi, \pi_2)$ . Note that this reduction process holds without assuming a preexisting Lagrangian formulation. Equations (7.14) together with their Hamiltonian structure can be obtained by Poisson reduction for cotangent bundles:  $T^*Q \rightarrow T^*Q/G$  (the so called Hamilton-Poincaré reduction [CMPR03]) applied here to the special case  $Q = TG$ .

We now check directly that equations (7.14) are equivalent to the  $2^{nd}$ -order Euler-Poincaré equations if the Hamiltonian (7.13) is associated to  $\ell$  by an invertible Legendre transform. The derivatives of the symmetry-reduced Hamiltonian  $h$  with respect to the momenta  $\pi_1$  and  $\pi_2$  imply formulas for the velocity and acceleration,

$$\frac{\delta h}{\delta \pi_1} = \xi, \quad \frac{\delta h}{\delta \pi_2} = \dot{\xi}, \quad (7.16)$$

so that the acceleration  $\dot{\xi}$  may be expressed as a function of the velocity  $\xi$  and the momenta  $(\pi_1, \pi_2)$ . The pair  $(\xi, \pi_2) \in T^*\mathfrak{g} \simeq \mathfrak{g} \times \mathfrak{g}^*$  obeys canonical Hamilton equations, so the derivatives of  $h$  in (7.13) with respect to velocity and acceleration imply the momentum relations,

$$\dot{\pi}_2 = -\frac{\delta h}{\delta \xi} = \frac{\delta \ell}{\delta \dot{\xi}} - \pi_1 \quad \text{and} \quad \frac{\delta h}{\delta \dot{\xi}} = 0 = \pi_2 - \frac{\delta \ell}{\delta \dot{\xi}}. \quad (7.17)$$

Solving these momentum relations for  $\pi_1$  and  $\pi_2$  in terms of derivatives of the reduced Lagrangian yields

$$\pi_1 = \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} \quad \text{and} \quad \pi_2 = \frac{\delta \ell}{\delta \dot{\xi}}. \quad (7.18)$$

The Lie-Poisson equation for  $\pi_1$ ,

$$\partial_t \pi_1 \pm \text{ad}_{\frac{\delta h}{\delta \pi_1}}^* \pi_1 = 0, \quad (7.19)$$

then implies the  $2^{\text{nd}}$ -order Euler-Poincaré equation,

$$(\partial_t \pm \text{ad}_{\xi}^*) \left( \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} \right) = 0.$$

### Example: Ostrogradsky-Lie-Poisson approach for geometric 2-splines

The Ostrogradsky reduced Hamiltonian (7.13) for geometric 2-splines is

$$h(\pi_1, \xi, \pi_2) = \frac{1}{2} \|\pi_2\|^2 - \langle \pi_2, \text{ad}_{\xi}^{\dagger} \xi \rangle + \langle \pi_1, \xi \rangle. \quad (7.20)$$

From this reduced Hamiltonian, the Poisson bracket (7.15) recovers the geometric 2-spline equations (3.20). For a bi-invariant metric  $\text{ad}_{\xi}^{\dagger} \xi = 0$  and these equations reduce to (3.21). In addition for  $SO(3)$  these equations produce the NHP equation (7.10).

**Third order.** Before going to the general case, it is instructive to quickly presenting the case of a third order  $G$ -invariant Lagrangian  $L : T^{(3)}G \rightarrow \mathbb{R}$ ,  $L = L(g, \dot{g}, \ddot{g}, \ddot{\ddot{g}})$  inducing the symmetry-reduced Lagrangian  $\ell : T^{(3)}G/G \simeq 3\mathfrak{g} \rightarrow \mathbb{R}$ ,  $\ell = \ell(\xi, \dot{\xi}, \ddot{\xi})$ . The Ostrogradsky momenta

$$p_1 = \frac{\partial L}{\partial \dot{g}} - \partial_t \frac{\partial L}{\partial \ddot{g}} + \partial_t^2 \frac{\partial L}{\partial \ddot{\ddot{g}}}, \quad p_2 = \frac{\partial L}{\partial \ddot{g}} - \partial_t \frac{\partial L}{\partial \ddot{\ddot{g}}}, \quad p_3 = \frac{\partial L}{\partial \ddot{\ddot{g}}},$$

produce the Legendre transform  $(g, \dot{g}, \dots, g^{(5)}) \in T^{(5)}G \mapsto (g, \dot{g}, \ddot{g}, p_1, p_2, p_3) \in T^*(T^{(2)}G)$ . The associated  $G$ -invariant Hamiltonian is obtained from the Legendre transformation,

$$H : T^*(T^{(2)}G) \rightarrow \mathbb{R}, \quad H(g, \dot{g}, \ddot{g}, p_1, p_2, p_3) := \langle p_1, \dot{g} \rangle + \langle p_2, \ddot{g} \rangle + \langle p_3, \ddot{\ddot{g}} \rangle - L(g, \dot{g}, \ddot{g}, \ddot{\ddot{g}}),$$

so that  $G$ -invariance of the Hamiltonian  $H$  yields the symmetry-reduced Hamiltonian  $h(\pi_1, \xi, \dot{\xi}, \pi_2, \pi_3)$ ,  $h : T^*(T^{(2)}G)/G \simeq \mathfrak{g}^* \times T^*(2\mathfrak{g}) \rightarrow \mathbb{R}$ . The reduced Hamiltonian  $h$  is related to the reduced Lagrangian  $\ell$  by the extended Legendre transformation

$$h(\pi_1, \xi, \dot{\xi}, \pi_2, \pi_3) = \langle \pi_1, \xi \rangle + \langle \pi_2, \dot{\xi} \rangle + \langle \pi_3, \ddot{\xi} \rangle - \ell(\xi, \dot{\xi}, \ddot{\xi}), \quad \frac{\delta \ell}{\delta \ddot{\xi}} = \pi_3. \quad (7.21)$$

The  $3^{\text{rd}}$ -order Ostrogradsky-Lie-Poisson system reads

$$\begin{cases} \partial_t \pi_1 \pm \text{ad}_{\frac{\delta h}{\delta \pi_1}}^* \pi_1 = 0, \\ \partial_t \xi = \frac{\delta h}{\delta \pi_2}, \quad \partial_t \pi_2 = -\frac{\delta h}{\delta \xi}, \quad \partial_t \dot{\xi} = \frac{\delta h}{\delta \pi_3}, \quad \partial_t \pi_3 = -\frac{\delta h}{\delta \dot{\xi}}, \end{cases} \quad (7.22)$$

with associated Poisson bracket

$$\{f, g\}(\pi_1, \xi, \dot{\xi}, \pi_2, \pi_3) = \{f, h\}_{\pm}(\pi_1) + \{f, h\}_{can}(\xi, \pi_2) + \{f, h\}_{can}(\dot{\xi}, \pi_3). \quad (7.23)$$

If the Hamiltonian (7.21) is associated to a Lagrangian  $\ell$  by Legendre transformation, we have

$$\dot{\xi} = \frac{\delta h}{\delta \pi_2}, \quad \dot{\pi}_2 = -\frac{\delta h}{\delta \xi} = -\pi_1 + \frac{\delta \ell}{\delta \xi}, \quad \ddot{\xi} = \frac{\delta h}{\delta \pi_3}, \quad \dot{\pi}_3 = -\frac{\delta h}{\delta \dot{\xi}} = -\pi_2 + \frac{\delta \ell}{\delta \dot{\xi}}, \quad \dot{\pi}_1 + \text{ad}^*_{\frac{\delta h}{\delta \pi_1}} \pi_1 = 0. \quad (7.24)$$

Consequently, we have

$$\pi_1 = \frac{\delta \ell}{\delta \xi} - \dot{\pi}_2 = \frac{\delta \ell}{\delta \xi} - \partial_t \left( \frac{\delta \ell}{\delta \dot{\xi}} - \dot{\pi}_3 \right) = \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} + \partial_t^2 \frac{\delta \ell}{\delta \ddot{\xi}},$$

and the last equation in (7.24) for the momentum map  $\pi_1$  implies by the 3rd-order Euler-Poincaré equation

$$(\partial_t \pm \text{ad}^*_{\xi}) \left( \frac{\delta \ell}{\delta \xi} - \partial_t \frac{\delta \ell}{\delta \dot{\xi}} + \partial_t^2 \frac{\delta \ell}{\delta \ddot{\xi}} \right) = 0.$$

**Higher-order and  $q$ -dependence.** The Ostrogradsky-Lie-Poisson approach generalizes to  $k^{\text{th}}$ -order as follows. For a  $G$ -invariant Lagrangian  $L : T^{(k)}G \rightarrow \mathbb{R}$ ,  $L = L(g, \dot{g}, \dots, g^{(k)})$ , the Ostrogradsky momenta define the Legendre transform as a map  $T^{(2k-1)}G \rightarrow T^*(T^{(k-1)}G)$  (see [dLR85]) and the associated Hamiltonian  $H : T^*(T^{(k-1)}G) \rightarrow \mathbb{R}$ ,  $H = H(g, \dot{g}, \dots, g^{(k-1)}, p_1, \dots, p_k)$  is given by

$$H(g, \dot{g}, \dots, g^{(k-1)}, p_1, \dots, p_k) := \sum_{j=1}^k \langle g^{(j)}, p_j \rangle - L(g, \dot{g}, \dots, g^{(k)}). \quad (7.25)$$

Extending the symmetry-reduced Ostrogradsky procedure outlined above to  $k^{\text{th}}$ -order then leads to the Ostrogradsky-Lie-Poisson equations

$$\begin{cases} \partial_t \pi_1 \pm \text{ad}^*_{\frac{\delta h}{\delta \pi_1}} \pi_1 = 0, \\ \partial_t \xi^{(j-2)} = \frac{\delta h}{\delta \pi_j}, \quad \partial_t \pi_j = -\frac{\delta h}{\delta \xi^{(j-2)}}, \quad j = 2, \dots, k, \end{cases} \quad (7.26)$$

whose Hamiltonian structure is

$$\{f, g\}(\pi_1, \xi, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k) = \{f, g\}_{\pm}(\pi_1) + \sum_{j=2}^k \{f, g\}_{can}(\xi^{(j-2)}, \pi_j). \quad (7.27)$$

This Poisson bracket produces the geometric  $k$ -spline equations from the corresponding reduced Hamiltonian; we do not carry out the details here.

The Ostrogradsky procedure for reduction by symmetry outlined above generalizes to allow  $q$ -dependence. Indeed, for a  $G$ -invariant Lagrangian  $L : (T^{(k)}G) \times Q \rightarrow \mathbb{R}$ , the previous steps may all be repeated with only slight changes, resulting in the reduced Poisson bracket (7.27), modified by adding the terms

$$\left\langle \mathbf{J} \left( \frac{\delta f}{\delta q} \right), \frac{\delta h}{\delta \pi_1} \right\rangle - \left\langle \mathbf{J} \left( \frac{\delta h}{\delta q} \right), \frac{\delta f}{\delta \pi_1} \right\rangle. \quad (7.28)$$

Thus, allowing  $q$ -dependence leaves the canonical equations invariant, but alters the  $\pi_1$ -equation so that it becomes

$$\partial_t \pi_1 \pm \text{ad}^*_{\frac{\delta h}{\delta \pi_1}} \pi_1 = -\mathbf{J} \left( \frac{\delta h}{\delta q} \right). \quad (7.29)$$

**From Clebsch to Ostrogradsky-Lie-Poisson.** Recall that, in the case  $k = 1$ , one passes from the canonical Clebsch formulation to the Lie-Poisson formulation (with Lie-Poisson bracket (7.2)) by using the momentum map, via the transformation

$$\alpha_q \in T^*Q \mapsto (\mu, q) \in \mathfrak{g}^* \times Q, \quad \mu = \mathbf{J}(\alpha_q).$$

Consider now the case  $k = 2$  with Lagrangian  $\ell = \ell(\xi, \dot{\xi}, q)$ . On the Clebsch side, the Hamiltonian is given by

$$H(\alpha_q, \xi, \pi) := \langle \pi, \dot{\xi} \rangle - \ell(\xi, \dot{\xi}, q) + \langle \mathbf{J}(\alpha_q), \xi \rangle, \quad \pi = \frac{\delta \ell}{\delta \dot{\xi}},$$

whereas the Ostrogradsky-Lie-Poisson Hamiltonian with  $q$ -dependence is defined by

$$h(\pi_1, \xi, \pi_2, q) := \langle \pi_1, \xi \rangle + \left\langle \pi_2, \dot{\xi} \right\rangle - \ell(\xi, \dot{\xi}, q), \quad \pi_2 = \frac{\delta \ell}{\delta \dot{\xi}}.$$

These definitions suggest that one can pass from the canonical Clebsch formulation on  $T^*(Q \times \mathfrak{g})$  to the Lie-Poisson-Ostrogradsky formulation on  $\mathfrak{g}^* \times T^*\mathfrak{g} \times Q$  by the Poisson map

$$(\alpha_q, \xi, \pi_2) \mapsto (\pi_1, \xi, \pi_2, q), \quad \pi_1 := \mathbf{J}(\alpha_q).$$

This is indeed the case, as one can check easily. Generalization to  $k > 2$  is now straightforward and we have the Poisson map

$$T^*(Q \times (k-1)\mathfrak{g}) \rightarrow \mathfrak{g}^* \times T^*(k-1)\mathfrak{g} \times Q,$$

$$(\alpha_q, \xi, \dot{\xi}, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k) \mapsto (\pi_1, \xi, \dot{\xi}, \dots, \xi^{(k-2)}, \pi_2, \dots, \pi_k, q), \quad \pi_1 := \mathbf{J}(\alpha_q),$$

that relates the canonical Hamilton equations on  $T^*(Q \times (k-1)\mathfrak{g})$  and the Ostrogradsky-Lie-Poisson equations (7.26).

## 8 Outlook and open problems

### 8.1 Brief summary and other potential directions

This paper has begun the application of symmetry-reduction tools to higher-order variational problems on Lie groups, culminating in an application of  $2^{nd}$ -order geometric splines to template matching on the sphere which was shown to be governed by a higher-order Euler-Poincaré equation on the dual Lie algebra of the Lie group  $SO(3)$ . The generality of this result was emphasized in Remark 5.2, on seeing that the higher-order Euler-Poincaré equation (5.15) had emerged once again as the optimality condition for template matching.

Various open problems not treated here seem to crowd together to present themselves. A few of these are:

- We have applied variational constraints to  $k$ -splines in Section 3.4. However, accommodating nonholonomic constraints would require additional developments of the theory.
- In Section 5 we presented higher-order methods that increased the smoothness in interpolating through a sequence of data points. In future work these methods will be compared to the shape spline model introduced in [TV10]. Some initial forays into the analysis of these problems were also presented in Section 5.2, but much remains to be done for these problems that have been treated here only formally.
- Extension of the basic theory presented here to allow for actions of Lie groups on Riemannian manifolds is also expected to have several interesting applications, particularly in image registration. For example, one could address higher-order Lie group invariant variational principles that include both curves on Lie groups and the actions of Lie groups on smooth manifolds, particularly on Riemannian manifolds. These Lie group actions on manifolds apply directly to the optimal control problems associated with large-deformation image registration in the *Large Deformation by Diffeomorphisms Metric Mapping* (LDDMM) framework via Pontryagin's maximum principle. Actions of Lie groups on Riemannian manifolds will be investigated in a subsequent treatment.

Doubtlessly, other opportunities for applying and extending this symmetry reduction approach for  $k$ -splines will present themselves in further applications.

## 8.2 An open problem: the slalom, or brachistochrone for splines

Let us formulate yet another example in slightly more detail. This is the brachistochrone version of the optimization problem treated here, for possible application, say, in a slalom race. Unlike the optimization problem, which seeks the path of least *cost*, a race would seek the path of least *time*. For example, the familiar *slalom* race involves dodging around a series of obstacles laid out on the course. The objective of the slalom racer in down-hill skiing, for example, is to pass through a series of gates as quickly as possible. The strategy in slalom racing is to stay close to the shortest-time path (or geodesic) between the gates, while also moderating the force exerted in turning to keep it below some threshold, lest the snow give way and the skier slides off the course. Thus, the ideal slalom path sought by an expert racer would hug the geodesic between the gates and make the series of turns passing through the gates with no skidding at all.

The strategy for achieving the optimal slalom has many potential applications in modern technology. For example:

- A charged-particle beam in an accelerator may be guided in its path by a series of quadrupole magnets that steer the beam to its target. The steering must be done as gently as possible, so as to minimize the transverse acceleration (seen as curvature in the path of the beam) that causes Bremsstrahlung and the consequent loss of energy in the beam.
- An underwater vehicle may be steered smoothly through narrow passageways in a sunken ship along a path that will take it quickly and efficiently to its objective, thereby avoiding collisions while minimizing fuel expenditure, time, etc.
- A car may be programmed to glide smoothly through a tight parallel-parking maneuver that ends in an elegant stop in the narrow space between two other cars along the curb.

- A vehicle may follow a program to roll as rapidly as possible along the terrain through a series of gates with its cameras mounted so that they continuously point toward an object above it that must keep in sight.

The slalom strategy that applies in all these examples seeks a path that minimizes the time for the distance travelled over a prescribed course, while also moderating the acceleration or force exerted along the path as it passes around a series of obstacles or through a series of gates laid out on the course. Designing such maneuvers requires optimization for least time, while also using cost functions that depend on both the velocity and acceleration of the motion. Moderation of higher-order accelerations such as *jerk* (rate of change of acceleration) may also be needed. As in the present paper, in solving optimal slalom problems that minimize the time taken to finish the course, one might expect to take advantage of continuous symmetries by investigating Lie group invariant variational problems for cost functions that are defined on  $k^{th}$ -order tangent spaces of Lie groups acting on smooth Riemannian manifolds. Investigations of invariant variational principles for slalom problems using the present group reduction and induced metric methods would be a promising direction. However, this direction seems even more challenging than the geometric splines for optimizing costs in trajectory planning on a Lie group treated here and it will be deferred to a later paper.

## Acknowledgements

The work of FGB was partially supported by a Swiss NSF postdoctoral fellowship. DDH is grateful for partial support by the Royal Society of London, Wolfson Scheme. TSR was partially supported by Swiss NSF grant 200020-126630. We also thank M. Bruveris and A. Trouvé for encouraging comments and insightful remarks during the course of this work.

## References

- [AFPA06] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, August 2006. PMID: 16788917.
- [Ash07] John Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38:95–113, 2007.
- [BC93] A. M. Bloch and P. E. Crouch. Nonholonomic and vakonomic control systems on Riemannian manifolds. *Fields Inst. Commun.*, 1, 1993.
- [BC96a] A. M. Bloch and P. E. Crouch. On the equivalence of higher order variational problems and optimal control problems. In *Proceedings of the 35th Conference on Decision and Control. Kobe, Japan*, pages 1648–1653, 1996.
- [BC96b] A. M. Bloch and P. E. Crouch. Optimal control and geodesic flows. *Systems & Control Lett.*, 28:65–72, 1996.
- [BFHL88] S. E. Bechtel, M. G. Forest, D. D. Holm, and K. J. Lin. One-dimensional closure models for three-dimensional incompressible viscoelastic free jets: von Kármán flow geometry and elliptical cross-section. *J. Fluid Mech.*, 196:241–262, 1988.

- [BGBHR10] M. Bruveris, F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu. The momentum map representation of images, 2010. Preprint available at <http://arxiv.org/abs/0912.2990>.
- [BK08] M. F. Beg and A. Khan. Representation of time-varying shapes in the large deformation diffeomorphic metric mapping framework. In *International Symposium of Biomedical Imaging*, 2008.
- [BMTY05] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *Int J. Comp. Vis.*, 61(2):139–157, 2005.
- [Bou71] N. Bourbaki. *Variétés différentielles et analytiques*. Fascicule de résultats / Paragraphes 8 à 15. Hermann, Paris, 1971.
- [CMPR03] H. Cendra, J. E. Marsden, S. Pekarsky, and T. S. Ratiu. Variational principles for Lie-Poisson and Hamilton-Poincaré equations. *Mosc. Math. J.*, 3(3):833–867, 2003.
- [CMR01] H. Cendra, J. E. Marsden, and T. S. Ratiu. Lagrangian reduction by stages. *Memoirs of the Amer. Math. Soc.*, 152(722):1–117, 2001.
- [CS95] P. E. Crouch and F. Silva Leite. The dynamic interpolation problem: On Riemannian manifolds, Lie groups, and symmetric spaces. *Journal of Dynamical and Control Systems*, 1(2):177–202, 1995.
- [CSC95] M. Camarinha, F. Silva Leite, and P. E. Crouch. Splines of class  $C^k$  on non-Euclidean spaces. *IMA Journal of Mathematical Control & Information*, 12:399–410, 1995.
- [dLR85] M. de León and P. R. Rodrigues. Formalismes Hamiltonien symplectique sur les fibrés tangents d’ordre supérieur. *C. R. Acad. Sc. Paris*, 301:455–458, 1985.
- [DPG<sup>+</sup>09] Stanley Durrleman, Xavier Pennec, Guido Gerig, Alain Trounev, and Nicholas Ayache. Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets, 2009.
- [FHT01] C. Foias, D. D. Holm, and E. S. Titi. The Navier-Stokes-alpha model of fluid turbulence. *Physica D*, 152:505–519, 2001. Preprint available at <http://arxiv.org/abs/nlin/0103037>.
- [GBHR10] F. Gay-Balmaz, D. D. Holm, and T. S. Ratiu. Geometric dynamics of optimization, 2010. Preprint available at <http://arxiv.org/abs/0912.2989>.
- [GBR09] F. Gay-Balmaz and T. S. Ratiu. The geometric structure of complex fluids. *Adv. in Appl. Math.*, 42:176–275, 2009.
- [GBR10] F. Gay-Balmaz and T. S. Ratiu. Clebsch optimal control formulation in mechanics, 2010. Preprint available at <http://arxiv.org/abs/0910.0874>.
- [GBT10] F. Gay-Balmaz and C. Tronci. Reduction theory for symmetry breaking with applications to nematic systems. *Physica D*, 239:1929–1947, 2010.



- [GK85] S. Gabriel and J. Kajiya. *Spline interpolation in curved space. State of the art in image synthesis*. SIGGRAPH 1985 course notes. ACM Press, New York, 1985.
- [HB04a] I. H. Hussein and A. M. Bloch. Dynamic interpolation on Riemannian manifolds: an application to interferometric imaging, 2004. Proceeding of the 2004 American Control Conference, Boston, MA.
- [HB04b] I. H. Hussein and A. M. Bloch. Optimal control on Riemannian manifolds with potential fields, 2004. 43rd IEEE Conference on Decision and Control, Paradise Island, Bahamas.
- [HM04] D. D. Holm and J. E. Marsden. Momentum maps and measure-valued solutions (peakons, filaments and sheets) for the EPDiff equation. *Progr. Math.*, 232:203–235, 2004. In *The Breadth of Symplectic and Poisson Geometry, A Festschrift for Alan Weinstein*.
- [HMR98] D. D. Holm, J. E. Marsden, and T. S. Ratiu. The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81, July 1998.
- [HTY09] D. D. Holm, A. Trouné, and L. Younes. The Euler-Poincaré theory of metamorphosis. *Quart. Appl. Math.*, 67:661–685, 2009. Preprint available at <http://arxiv.org/abs/0806.0870>.
- [JOP<sup>+</sup>] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.
- [KM97] A. Kriegl and P. Michor. *The Convenient Setting of Global Analysis*, volume 53 of *Surveys and Monographs*. American Mathematical Society, 1997.
- [Mic08] P. W. Michor. *Topics in Differential Geometry*, volume 93 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2008.
- [MSK10] L. Machado, F. Silva Leite, and K. Krakowski. Higher-order smoothing splines versus least squares problems on riemannian manifolds. *J. Dyn. and Control Syst.*, 16:121–148, 2010.
- [MTY02] Michael I. Miller, Alain Trouné, and Laurent Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405, 2002.
- [MW83] J. E. Marsden and A. Weinstein. Coadjoint orbits, vortices, and Clebsch variables for incompressible fluids. *Physica D*, 7:305–323, 1983.
- [MY01] M.I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41(1-2):61–84, 2001.
- [NHP89] L. Noakes, G. Heinzinger, and B. Paden. Cubic splines on curved spaces. *IMA Journal of Mathematical Control & Information*, 6:465–473, 1989.
- [Noa03] L. Noakes. Null cubics and Lie quadratics. *J. Math. Phys.*, 44:1436–1448, 2003.

- [Noa04] L. Noakes. Non-null Lie quadratics in E3. *J. Math. Phys.*, 45:4334–4351, 2004.
- [Noa06] L. Noakes. Duality and Riemannian cubics. *Adv. Comput. Math.*, 25:195–209, 2006.
- [Pop07] T. Popiel. Higher order geodesics in Lie groups. *Math. Control Signals Syst.*, 19:235–253, 2007.
- [TV10] Alain Trouvé and François-Xavier Vialard. Shape splines and stochastic shape evolutions: A second order point of view, 2010.
- [TY05] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM J. Math. Anal.*, 37(1):17–59, 2005.
- [Via09] F.-X. Vialard. *Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework : From the Discontinuous Image Matching Problem to a Stochastic Growth Model*. PhD thesis, Ecole Normale Supérieure de Cachan, 2009. <http://tel.archives-ouvertes.fr/tel-00400379/fr/>.
- [You10] L. Younes. *Shapes and Diffeomorphisms*. Applied Mathematical Sciences. Springer, 1st edition, September 2010.